

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:32:52 ; Search time 168.5 Seconds
(without alignments)
4483.584 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atgcgcgtcccgcttgggaa.....cagctcagcaagaagcttaa 1053

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO spool D/US10089688/runat 10112004 152246 4835/app query.fasta_1.1223
-DB=A Geneseq 23Sep04 -OPMT=fastan -SUFFIX=rag -MINMATCH=0 1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN 1 1 224 @runat 10112004 152246 4835 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1781	98.8	350	AAB20387	Aab20387 Human acu
2	1635.5	90.7	341	AAB48970	Aab48970 Human ANI
3	1635.5	90.7	341	ADJ69105	Adj69105 Human hea
4	1635.5	90.7	496	AAB10858	Aae10858 Gal4-huma
5	1635.5	90.7	552	AAB10859	Aae10859 LexA-huma
6	1614.5	89.5	341	AAY94248	Aay94248 Mouse cal
7	1597.5	88.6	354	ABG23844	Abg23844 Novel hum
8	1297.5	72.0	337	AAY94247	Aay94247 Human cal
9	1297.5	72.0	337	AAB82090	Aab82090 Human acu
10	1297.5	72.0	337	AAM39078	Aam39078 Human pol

11	1153.5	64.0	289	4	AAB94139	Aab94139 Human pro
12	1123	62.3	339	4	ABB60392	Abb60392 Drosophil
13	1117	62.0	339	3	AAI94249	Aay94249 Drosophil
14	1054.5	58.5	337	3	AAI94250	Aay94250 C. elega
15	888.5	49.3	237	4	AAM40864	Aam40864 Human pol
16	705.5	39.1	343	3	AAG45273	Aag45273 Arabidops
17	674	37.4	326	3	AAG51053	Aag51053 Arabidops
18	674	37.4	400	3	AAG51052	Aag51052 Arabidops
19	674	37.4	504	3	AAG51051	Aag51051 Arabidops
20	673.5	37.4	300	3	AAG23886	Aag23886 Arabidops
21	669.5	37.1	345	3	AAG05089	Aag05089 Arabidops
22	666.5	37.0	300	3	AAG45274	Aag45274 Arabidops
23	666.5	37.0	300	3	AAG30714	Aag30714 Arabidops
24	666.5	37.0	300	3	AAG30713	Aag30713 Arabidops
25	658.5	36.5	285	3	AAG23887	Aag23887 Arabidops
26	635	35.2	320	3	AAG05090	Aag05090 Arabidops
27	523.5	29.0	213	3	AAG23888	Aag23888 Arabidops
28	518.5	28.8	213	3	AAG30715	Aag30715 Arabidops
29	518.5	28.8	213	3	AAG45275	Aag45275 Arabidops
30	465.5	25.8	154	3	AAG41151	Aag41151 Zea mays
31	460	25.5	197	3	AAG05091	Aag05091 Arabidops
32	448.5	24.9	148	3	AAG41152	Aag41152 Zea mays
33	430.5	23.9	139	3	AAG41153	Aag41153 Zea mays
34	276	15.3	639	4	ABG25372	Abg25372 Novel hum
35	271	15.0	383	4	ABG23843	Abg23843 Novel hum
36	269	14.9	236	5	ABP02921	Abp02921 Human ORF
37	209	11.6	135	5	ABP34081	Abp34081 Human ORF
38	173	9.6	75	4	ABG25370	Abg25370 Novel hum
39	127.5	7.1	121	3	AAG33914	Aag33914 Arabidops
40	126	7.0	1182	6	ABU25521	Abu25521 Protein e
41	116.5	6.5	888	5	AAB83013	Aau83013 Human hom
42	116.5	6.5	972	6	ABU19394	Abj19394 Human int
43	113	6.3	541	7	ADH87535	Ade87535 Enterococ
44	113	6.3	2184	4	AAE00425	Aae00425 P. faicip
45	111.5	6.2	790	7	ADJ70372	Adj70372 Human hea

ALIGNMENTS

RESULT 1
AAB20387

ID AAB20387 standard; protein; 350 AA.

XX AC AAB20387;

XX DT 11-JUN-2001 (first entry)

XX DE Human acute neuronal induced calcium binding protein ANIC-BP-1B.

XX KW Acute neuronal induced calcium binding protein; ANIC-BP-1B;
KW spice variant; human; stroke; head trauma; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; spinal cord injury;
KW cerebroprotective; antiparkinsonian; nootropic; neuroprotective; therapy;
KW diagnosis; vaccine.

XX OS Homo sapiens.

XX PN WO200125423-A1.

XX PD 12-APR-2001.

XX PF 28-SEP-2000; 2000WO-EP009475.

XX PR 04-OCT-1999; 99EP-00119113.

XX PA (MERE) MERCK PATENT GMBH.

XX PI Duecker K, Den Daas I;

XX DR WPI; 2001-266306/27.

XX DR N-PSDB; AAF30688.

PT Novel human acute neuronal induced calcium-binding protein like protein

PT splice variant, useful for treating stroke, acute head trauma,
 PT Parkinson's disease, Alzheimer's disease multiple sclerosis, spinal cord
 PT injury.

XX
 PS
 XX Claim 2; Page 44-45; 49pp; English.

CC The present sequence is that of a novel human acute neuronal induced
 CC calcium binding protein-like protein splice variant, ANIC-BP-1B. The
 CC protein shows homology to other members of the calcium binding protein
 CC family, including ANIC-BP, a protein discovered by mRNA differential
 CC display that is upregulated in a rat model of head trauma. ANIC-BP and
 CC ANIC-BP-1B differ in their C-terminal portions. The variant protein could
 CC serve as a novel drug target. The invention provides ANIC-BP-1B
 CC polynucleotides (see AAF30688) and polypeptides, expression vectors, host
 CC cells and antibodies, as well as methods for producing the protein and
 CC for treating or preventing disorders associated with expression of the
 CC protein by inhibiting or activating the action of ANIC-BP-1B. Diseases
 CC that may be treated include stroke and acute head trauma, Parkinson's
 CC disease, Alzheimer's disease, multiple sclerosis and spinal cord injury.
 CC The polynucleotides and polypeptides can also be used in diagnostic
 CC assays and in vaccines, and to identify agonists and antagonists useful
 CC for treating conditions associated with ANIC-BP-1B imbalance

XX
 SQ Sequence 350 AA;

Alignment Scores:

Pred. No.: 2.53e-178 Length: 350
 Score: 1781.00 Matches: 350
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.78% Indels: 0
 DB: 4 Gaps: 0

US-10-089-688-1 (1-1053) x AAB20387 (1-350)

Qy 1 ATGCGGTCCCGTGGAGCTCTCAAAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
 Db 1 MetProPheProPheGlySerHisLysSerProAlaAspIleValLysAsnLeuLys 20
 Qy 61 CAGAGCAGTGGCTGTTCTCGAAAGCAAGACATTTCTGATAAAAGCAGAAAGCGTACA 120
 Db 21 GluSerMetAlaValLeuLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
 Qy 121 CAGAAGTTTCAAAATCTGGTCCCATGAAAGAAATCTGTATGGCACAATAAGAAAA 180
 Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
 Qy 181 GAGCTCAGCAGAGCAGTAGCTCAACTCTCTCAAGAACTCTATAATAGTGGCTCCTT 240
 Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
 Qy 241 AGCACCTTGGTAGCTGATTATCAGCTCAATTCAGCTTTTGAAGGCAAAAGAGCGTGGCTCAA 300
 Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
 Qy 301 ATTTTCAACAATTTCTCAGAGACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
 Db 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
 Qy 361 TGCACCCCAACAGAAATTTTGTTCATGTTATTGAAGGGTATGAATCTCCAGAAATAGCT 420
 Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
 Qy 421 CTAATTTGTGAATAATTTGAAGAATGCGATCAGACATGAACCACTTGCAGAAATCAATT 480
 Db 141 LeuAsnGlyGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
 Qy 481 TTGTGGTGGACAGTTTATGATTTCTTCAGATATGTCGAAATGTCGAAATTCACATTTCACATA 540
 Db 161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
 Qy 541 GCTTCAGATGCTTTGGCCACATTTCAAGGATTTTACTTACAAGACATAAATTTGCTCAGTGCA 600

Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 Qy 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATCAGAGTACTTCAATTCATCA 660
 Db 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 Qy 661 GAAAAATTATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAAACTACTACTAGATAGA 720
 Db 221 GLUASHTYRVALTHRYLSARGGLNSERLEULYSLEULYSGLYGLULLEULYEAAPARG 240
 Qy 721 CACAACCTTCAAAATTATGACAAATAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
 Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 Qy 781 AACCTGCTGGCAGACAAAAAGTCCCAACATCCAGTTTTCAGGCGCTTTCACGTTTTTAAGGTG 840
 Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 Qy 841 TTTGTAGCCAATCTTAACAGACGCGAGCCCATCTAGACATCTCTCAAGAAACCGAGCC 900
 Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLysAsnGlnAla 300
 Qy 901 AAACTCATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGAGGACGGATTGTATGAGCAGTCC 960
 Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrAspCysMetSerSerSer 320
 Qy 961 GTACCGAGCAGCAATTTCCCGGTTCGATTACGGGTTAAACCGCGGAGCGTGGATCAGG 1020
 Db 321 ValProThrThrAsnSerArgValAspLeuArgValLysProArgThrArgGlyIleArg 340
 Qy 1021 GATTTGAAGACGACGCTCAGCAAGAGCT 1050
 Db 341 AspLeuLysArgProAlaGlnGlnGluAla 350

RESULT 2

AAB48970
 ID AAB48970 standard; protein; 341 AA.
 XX
 AC AAB48970;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Human ANIC-BP (acute neuronal induced calcium-binding protein).
 XX
 KW Human; acute neuronal induced calcium-binding protein; ANIC-BP;
 KW Mo25 homologue; Hyma homologue; drug screening; stroke;
 KW acute head trauma; multiple sclerosis; spinal cord injury; vaccine;
 KW cerebroprotective; neuroprotective.
 XX
 OS Homo sapiens.
 XX
 FN WO200078947-A1.
 XX
 PD 28-DEC-2000.
 XX
 PF 14-JUN-2000; 2000WO-EP005457.
 XX
 PR 22-JUN-1999; 99EP-00112024.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Den Daas I, Fischer V, Seyfried C, Von Melchner L;
 XX
 DR WPI; 2001-102721/11.
 XX
 DR N-PSDB; AAC91772.
 XX
 PT Novel acute neuronal induced calcium binding protein, useful for treating
 PT acute head trauma, stroke, multiple sclerosis and spinal cord injury.
 XX
 PS Claim 2; Page 37; 50pp; English.
 XX
 CC The invention relates to human acute neuronal induced calcium-binding
 CC protein (ANIC-BP) and to nucleic acid encoding it. The invention also

CC relates to expression systems and recombinant host cells comprising ANIC-
CC BP DNA, the recombinant production of ANIC-BP, antibodies specific for
CC ANIC-BP, fusion proteins comprising ANIC-BP and an immunoglobulin Fc
CC region, and methods of screening for modulators of ANIC-BP function. ANIC
CC -BP has homology and structural similarity to HymA and Mo25 proteins.
CC ANIC-BP proteins and nucleotides are useful for treating stroke and acute
CC head trauma, multiple sclerosis and spinal cord injury. ANIC-BP proteins
CC are useful in screening assays, for identifying membrane bound or soluble
CC receptors, and also in vaccines. ANIC-BP nucleotides are useful as
CC diagnostic reagents, as tools for tissue expression studies, for
CC chromosome localisation studies, as genetic vaccines, and in the
CC generation of transgenic animals. The present sequence represents human
CC ANIC-BP
XX
SQ Sequence 341 AA;

Alignment Scores:
Pred. No.: 5,43e-163 Length: 341
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAB48970 (1-341)

QY 1 ATCCGCTCCGTTGGAGTCTCAAAATCTCCAGACAGATTTGGAAGATCTGAAG 60
DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGCTGCTTCGGAAGACAGACATTTCTGATTAAGAGCAGAAAGGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAATCTGGTTCATGCAAGAAATCTGTATGGCACAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
QY 181 GAGCTCAGACAGACAGTAGCTCAATCTGCTCAAGAACTCTATATATGCGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysLysAsnSerGlyLeuLeu 80
QY 241 AGACCTGCTAGCTGATTTACAGCTCATTTGACTTTGAGGGCAAAAAGACGCTGCTCA 300
DB 81 SerThrLeuValAlaAspLeuGlnLeuLeuAlaAspPheGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACATATTTCTCAGACAGACAAATTTGGTACGAGAACTCTCTCTGTGAATACATC 360
DB 101 IlePheAsnAsnIleLeuAArgGlnIleGlyThrArgThrProThrValGluLysIle 120
QY 361 TGCACCCAAACAGATATTTTGTTCATGTTATTTGAAGGGTATGAATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyLysGlyThrGluSerProGluIleAla 140
QY 421 CTAATTTGCGAATATGTTAAGAGATGTCATCAGACATCAACCACTTGCAAAAATCAAT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
QY 481 TTGTGTCGGAACAGTTTATGATTTTCTCAGATATGTCGAAATGTCAACATTTGACATA 540
DB 161 LeuTrpSerGluGlnPheThrAspPhePheArgTrpValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGATTTGCCACATTTCAAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTGCAGACAGATTTATGATGATTTTTCAGTGAATATGAGAGTACTTCAATCA 660
DB 201 GluPheLeuGluGlnHisLysTrpAspArgPhePheSerGluTrpGluLysLeuHisSer 220
QY 661 GAAATTTATGTGACAAAGACAGTCACTGAAGCTTCTCGGTCAACTACTACTAGATAGA 720
DB 221 GluAsnTrpValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240

QY 721 CACAACTTCAATATTGACAAATAATACATCAGTAAACCTGAGAACCTCAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACTCTGCTGAGACAAACCTCCCAACATCCAGCTTTTCAGGCTTTTCAGTTTAAAGTG 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCAAATCTTAACAGAGCAGCCCATCTCAGACATCTCTCAAGAACAGAGGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
QY 901 AACTCATAGATTTCTCAGCAAGTTTTCAGACAGACAGAGCGAGATTTGATGAGCAGTTCC 960
DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCCGAGCAGGAATTCCTCCGGTTCGATTACGGCTTAACCGCGGACGCGTGGGATCAGG 1020
DB 316 AspGluGlnPheAsnAspGluLysThrTyLeuValLysGln-----IleArg 331
QY 1021 GATTTCAGAGAGACAGCTCAGCAAGAGCT 1050
DB 332 AspLeuLysArgProAlaGlnGlnGluAla 341

RESULT 3

ADJ69105
ID ADJ69105 standard; protein; 341 AA.

AC ADJ69105;

DT 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SeqID911.

DE mitochondria; human; screening assay; diabetes mellitus;

XX Huntington's disease; osteoarthritis;

KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;

KW osteopathic; ophthalmological; cytostatic.

OS Homo sapiens.

XX WO2003087768-A2.

PN 23-OCT-2003.

PD 04-APR-2003; 2003WO-US010870.

XX 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-0389987P.

PR 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

PA (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

PI Warnock DE;

XX WPI; 2003-845369/78.

DR Identifying a mitochondrial target for drug screening assays and for

XX treating diseases associated with altered mitochondrial function,

PT comprises detecting a modified polypeptide in a sample and correlating

PT with the disease.

XX Claim 1; SEQ ID NO 911; 180pp; English.

PS This invention relates to novel mitochondrial targets that can be used

CC for therapeutic intervention in treating a disease associated with

CC altered mitochondrial function. Specifically, it refers to a method for

CC	identifying proteins of the human heart mitochondrial proteome that are
CC	useful for drug screening assays, as well as therapeutic targets. The
CC	present invention describes a method for identifying such proteins that
CC	can be used in the treatment of various diseases associated with altered
CC	mitochondrial function including diabetes mellitus, Huntington's disease,
CC	osteochondritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC	ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC	compositions have neuroprotective, nootropic, antidiabetic,
CC	anticonvulsant, antiarrhythmic, osteoprotic, ophthalmological and
CC	cytostatic activities. This polypeptide sequence is a human heart
CC	mitochondrial protein of the invention.
XX	
SQ	Sequence 341 AA;
Alignment Scores:	
Pred. No.:	5 43e-163 Length: 341
Score:	1635.50 Matches: 329
Percent Similarity:	94.29% Conservative: 1
Best Local Similarity:	94.00% Mismatches: 11
Query Match:	90.71% Indels: 9
DB:	7 Gaps: 2
US-10-089-688-1 (1-1053) x ADJ69105 (1-341)	
QY	1 ATGCGGTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGCATTTGTGAAGATCTGAAG 60
DB	1 MetProheProheGlyysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY	61 GAGAGCATGCTGTTCTGGAAGACAGACATTTCTGATTAATAAAGCAGAAAGGCTACA 120
DB	21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY	121 GAAGAGTTCCTCAAAATCTGTTGCCATCAAGAAATTTCTGTATGCGCACAAATGAAAA 180
DB	41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
QY	181 GAGCCTCAGACAGACAGTAGCTCAACTGCTGCTCAAGACTCTATATATAGTGGCTCTT 240
DB	61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysLeuTyAsnSerGlyLeuLeu 80
QY	241 AGCACCTGTTAGTGTATTTACAGTCTCATTTGAGGGCAAAAGAGCTGGCTCAA 300
DB	81 SerThrLeuValAlaAspLeuGlnLeuLysAspPheGluGlyLysAspValAlaGln 100
QY	301 ATTTTCAACATATTTCTCAGAAACAAATTTGTCAGAGAACTCTCTACTGTTGAATATCATC 360
DB	101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyIle 120
QY	361 TGCACCCACAGATATTTTGTGTTATGTTATGAAAGGTATGAATCTCCAGAAATAGCT 420
DB	121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyGluSerProGluIleAla 140
QY	421 CTAAATTTGCAATATGTTTAAAGAAATGTCATGACATGAACCTTGCAAAATCATT 480
DB	141 LeuAsnCysGlyIleMetLeuArgGlnCysIleArgHisGluProLeuAlaLysIleIle 160
QY	481 TTGTCGTCGGAACAGCTTTTATGATTTCTTCAGATATGTCGAATGTCGAATTTGACATA 540
DB	161 LeuTrpSerGluGlnPheTyAspPhePheArgTyValGluMetSerThrPheAspIle 180
QY	541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATTAATTCCTCATGTCA 600
DB	181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY	601 GAATTTTGGACACATTTATGATGATTTTTCAGTGAATATGAGAGTACTTCTCATTTCA 660
DB	201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyGluLysLeuLeuHisSer 220
QY	661 GAAATTTATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGA 720
DB	221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240
QY	721 CACAACTTCACAAATATATGACAAAATATACATGATAAACCCTGAGAACCTCAAAATATATGATG 780
DB	241 HisAsnPheThrIleMetThrLysTyIleSerLysProGluAsnLeuLysLeuMetMet 260
QY	781 AACCTGCTCGGAGACAAAGTCCGAAATCCAGTTTGGCCCTTTTCAGCTTTTAAAGGTG 840
DB	261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY	841 TTTGTAGCCAACTTAAACAAGACGAGCCATCTCTAGACATCTCTCAAGAACCCAGGCGC 900
DB	281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
QY	901 AAACCTCATAGACTCTCTCAGCAAGTTTCAGAACGACGAGGATTTGATGACAGTTCC 960
DB	301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu 315
QY	961 GTACCGACGACGAATTCCTCCGGGTGCGATTACCGCTTAACCCGGGCGGTGGATCAGG 1020
DB	316 AspGluGlnPheAsnAspGluLysThrTyLeuValLysGln 1050
QY	1021 GATTTGAAGACAGACCTCAGCAAGAGCT 1050
DB	332 AspLeuLysArgProAlaGlnGlnGluAla 341
RESULT 4	
AAE10858	
ID	AAE10858 standard; protein; 496 AA.
XX	
AC	AAE10858;
XX	
DT	18-DEC-2001 (first entry)
XX	
DE	Gal4-human ANIC-BP-1 fusion protein.
XX	
KW	Human; acute neuronal induced calcium binding protein type 1 ligand;
KW	ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW	Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW	gene therapy; fusion protein; Gal4 protein.
XX	
OS	Homo sapiens.
OS	Unidentified.
OS	Chimeric.
XX	
PN	WO200170771-A2.
XX	
PD	27-SEP-2001.
XX	
PF	20-MAR-2001; 2001WO-EP003149.
XX	
PR	21-MAR-2000; 2000EP-00106110.
XX	
PA	(MERE) MERCK PATENT GMBH.
XX	
PI	Den Daas I, Duecker K, Hock B;
XX	
DR	WPI; 2001-607519/69.
XX	
PT	Novel acute neuronal induced calcium binding protein type 1 ligand
PT	polypeptides, useful in the treatment of stroke, head trauma, multiple
PT	sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
PT	injury.
XX	
PS	Disclosure; Page 42-44; 46pp; English.
XX	
CC	The invention relates to human acute neuronal induced calcium binding
CC	protein type 1 (ANIC-BP-1) ligand polypeptides and polymucleotides.
CC	Sequences of the invention are useful for treating human diseases
CC	including stroke, head trauma, multiple sclerosis, Parkinson's disease,
CC	Alzheimer's disease and spinal cord injury. They are also useful as
CC	vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
CC	soluble receptors. Polynucleotides of the invention are useful as
CC	diagnostic reagents, for chromosome localization studies, and as valuable
CC	tools for tissue expression studies. They are also useful in gene

CC therapy. The present sequence is Gal4-human ANIC-BP-1 fusion protein
CC comprising the Gal4 protein and a C-terminally linked human ANIC-BP-1
CC protein

XX SQ Sequence 496 AA;

Alignment Scores:

Pred. No.: 6.26e-163 Length: 496
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AAE10858 (1-496)

QY 1 ATGCGGTTCCCGTTGGGAGCTCTACAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
DB |||||
QY 156 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 175
DB |||||
QY 61 GAGACATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAAGAGCAAGAAAGCGTACA 120
DB |||||
QY 176 GluSerMetAlaValLeuGluLysGluAspIleSerAspLysAlaGluLysAlaThr 195
DB |||||
QY 121 GAAGAAGTTTCAAAATCTGGTGGCCNTGAAAGAAATCTGTATGGCAGCAAAATGAAAA 180
DB |||||
QY 196 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 215
DB |||||
QY 181 GAGCTCAGACAGACAGTAGCTCACTCTCAAGACTCTAATAATAGTGGGCTCCTT 240
DB |||||
QY 216 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysSerGlyLeuLeu 235
DB |||||
QY 241 AGCACCTGGTAGCTGATTTACAGCTCATTCAGCTTTGAGGCAAAAGAGCGTGGCTCAA 300
DB |||||
QY 236 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 255
DB |||||
QY 301 ATTTTCAACAATATCTCAGAAAGCAAAATGGTAGAGAACTCTACTGTGTAATCATC 360
DB |||||
QY 256 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluLys 275
DB |||||
QY 361 TGACCCCAACAGATATTTTCTCATGTTATTTGAAGGATGATGAATCTCCAGAAATAGCT 420
DB |||||
QY 276 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyThrGluSerProGluIleAla 295
DB |||||
QY 421 CTAAATTTGTGAATTAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCAT 480
DB |||||
QY 296 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 315
DB |||||
QY 481 TTGTGGTCGGAACAGTTTTATGATTTCTCAGATATGTCGAAATGTCACATTTGACATA 540
DB |||||
QY 316 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 335
DB |||||
QY 541 GCTTCAGATGATTTGCAATTCAGATTTCAAGATTTACTTACAGACATAAATGCTCAGTGCA 600
DB |||||
QY 336 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 355
DB |||||
QY 601 GAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATAGAGAGTACTTCAATTC 660
DB |||||
QY 356 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 375
DB |||||
QY 661 GAAATTTATGTGACAAAGACAGTCACTAGCTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
DB |||||
QY 376 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 395
DB |||||
QY 721 CACAACTTCAAAATTTGACAAATATCATCATCAAGTAACTGAGACCTCAATTAATCATG 780
DB |||||
QY 396 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 415
DB |||||
QY 781 AACCTGCTGCGAGACAAAGTCGAAACATCCAGTTTGGAGCTTTCAGGCTTTTCAGTTTAAAGTG 840
DB |||||
QY 416 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 435
DB |||||
QY 841 TTTGTAGCCAAATCTTAACAGACGACGCCCATCTAGACATCTCTCTCAAGAACCGGCC 900

DB |||||
QY 436 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLysAsnGlnAla 455
DB |||||
QY 901 AAATCTCATAGATTCTCTCAGCAAGTTTTCAGAACGACAGGAGGATTTGTATGAGCAGTTCC 960
DB |||||
QY 456 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 470
QY 961 GTACCGACGACGAATTTCCCGGTGCGATTAGCGGTTHAAACGCGGACGCGTGGGATCAGG 1020
DB |||||
QY 471 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 486
QY 1021 GATTGGAAGACACAGCTCAGCAAGAGCT 1050
DB |||||
QY 487 AspLeuLysArgProAlaGlnGlnGluAla 496

RESULT 5
AAE10859
ID AAE10859 standard; protein; 552 AA.
XX
AC AAE10859;
XX
DT 18-DEC-2001 (first entry)
XX
DE LexA-human ANIC-BP-1 fusion protein.
XX
KW Human; acute neuronal induced calcium binding protein type 1 ligand;
KW ANIC-BP-1; human disease; stroke; head trauma; multiple sclerosis;
KW Parkinson's disease; Alzheimer's disease; spinal cord injury; vaccine;
KW gene therapy; fusion protein; LexA protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PH Key Location/Qualifiers
FT Region 1..202
FT /note= "LexA protein"
FT Region 203..552
FT /note= "Human ANIC-BP-1 protein"
XX
XX WO200170771-A2.
XX
XX 27-SEP-2001.
XX
XX 20-MAR-2001; 2001WO-EP003149.
XX
XX 21-MAR-2000; 2000EP-00106110.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Den Daas I, Duecker K, Hock B;
XX
XX WPI; 2001-607519/69.
XX
XX Novel acute neuronal induced calcium binding protein type 1 ligand
XX polypeptides, useful in the treatment of stroke, head trauma, multiple
XX sclerosis, Parkinson's disease, Alzheimer's disease and spinal cord
XX injury.
XX
XX Disclosure; Page 44-46; 46pp; English.
XX
XX The invention relates to human acute neuronal induced calcium binding
XX protein type 1 (ANIC-BP-1) ligand polypeptides and polynucleotides.
XX Sequences of the invention are useful for treating human diseases
XX including stroke, head trauma, multiple sclerosis, Parkinson's disease,
XX Alzheimer's disease and spinal cord injury. They are also useful as
XX vaccines. ANIC-BP-1 ligands are useful for identifying membrane bound
XX soluble receptors. Polynucleotides of the invention are useful as
XX diagnostic reagents, for chromosome localization studies, and as valuable
XX tools for tissue expression studies. They are also useful in gene
XX therapy. The present sequence is LexA-human ANIC-BP-1 fusion protein
XX comprising the LexA protein and a C-terminally linked human ANIC-BP-1
XX protein

```
XX SQ Sequence 552 AA;
Alignment Scores:
Pred. No.: 6,52e-163 Length: 552
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservativity: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 4 Gaps: 2

US-10-089-688-1 (1-1053) x AA010859 (1-552)
QY 1 ATGCGGTTCCGGTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 212 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 231
QY 61 GAGAGCATGGCTGTTCTGGAAGAACAGACATTTCTGATAAAAGCAGAAAGCGCTACA 120
DB 232 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 251
QY 121 GAAGAAGTTTCCAAAATCTGTTGCCATGAAGAAATTTCTGTATGGCACAATAAGAAA 180
DB 252 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 271
QY 181 GAGCCTCAGACAGACAGTAGTCACTTGTCTCAAGAACTCTATAATAGTGGCTCCTT 240
DB 272 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 291
QY 241 AGCACCTGGTAGCTGATTTACAGCTCATTTGAGCTTTGAGGGCAAAAAGAGCTGGCTCAA 300
DB 292 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 311
QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTCGTACGAGAACTCTACTGTTGAATACATC 360
DB 312 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 331
QY 361 TGCACCCACAGAAATTTGTTGTTATGTTTGAAGGGTATGAATCTCCAGAAATAGCT 420
DB 332 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 351
QY 421 CTAAATTTGGATATTTAGAGATGCAATGACATGACATGAACCACTTGGCAAAATCATTT 480
DB 352 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 371
QY 481 TTGTGGTCGGAACAGTTTTATGATTCTTCAGATATGTCGAAATGTCACATTTGACATA 540
DB 372 LeuIlePheSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 391
QY 541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAAGACATATAAATTCCTCAGTGCA 600
DB 392 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 411
QY 601 GAATTTTGGACACGATATGATAGATTTTTCAGTGAATAGAGAGTTACTTCAATCA 660
DB 412 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 431
QY 661 GAAATATTATGTACAAAAGACAGTCACCTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
DB 432 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 451
QY 721 CACAACTTTCACAAATATGACAAATATCATCATGTAACCTGAGAACCTCAAAATTAATGATG 780
DB 452 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 471
QY 781 AACCTGCTCGGACAAAAGTCGCAACATCCAGTTTGGAGCTTTCACGTTTAAAGTG 840
DB 472 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 491
QY 841 TTTGTAGCAATCTTAAACAGACGAGCCCATCTCTAGACATCTCTCAAGAACACAGGCC 900
DB 492 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 511
```

```
901 AAACATATAGAGTTCTCTCAGCAAGTTTCAGAACAGACAGGAGGATTTGATGACAGATTCC 960
DB 512 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 526
QY 961 GTACCGACGACGAATTCCTCCGGTTCGATTACGGGTTAAACCGCGGACGCGTGGGATCAGG 1020
DB 527 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 542
QY 1021 GATTTGAAGACAGCAGCTCAGCAAGAGAGCT 1050
DB 543 AspleuLysArgProAlaGlnGlnGluAla 552

RESULT 6
AAAY94248
ID AAY94248 standard; protein; 341 AA.
AC AAY94248;
XX 10-AUG-2000 (first entry)
DT Mouse calcium binding protein MO25.
DE
KW Mouse; calcium binding protein; cancer; inflammation; MO25; CBP;
KW reproductive disorder; autoimmune disorder; developmental disorder;
KW seizure disorder; immune disorder; infection.
OS Mus sp.
PN WO200029580-A1.
XX 25-MAY-2000.
PD
PF 12-NOV-1999; 99WO-US027027.
XX 13-NOV-1998; 98US-00190965.
PR
XX (INCY-) INCYTE PHARM INC.
PA
XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;
XX WPI; 2000-387793/33.
DR
XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
XX diagnosis, prevention and treatment of cancers, immune, developmental or
XX reproductive disorders.
PS Disclosure; Page 66-67; 72pp; English.
CC
CC The present sequence is the mouse calcium binding protein MO25. It was
CC used in a sequence alignment to identify human calcium binding protein
CC hCBP. The hCBP protein and the gene encoding it are useful for the
CC diagnosis and treatment of the following types of disorder: cancers (such
CC as adenocarcinomas), reproductive disorders (such as infertility,
CC ovulatory defects, endometriosis, disruptions of the oestrus and
CC menstrual cycles, polycystic ovary syndrome and ovarian
CC hyperstimulation), autoimmune disorders (such as benign prostatic
CC hyperplasia and prostatitis), developmental disorders (such as Cushing's
CC syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
CC neuropathies, seizure disorders, immune disorders (such as AIDS,
CC allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's
CC disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
CC rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
CC colitis), and viral, bacterial, fungal, parasitic, protozoal and
CC helminthic infections
XX
SQ Sequence 341 AA;

Alignment Scores:
Pred. No.: 8,89e-161 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservativity: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55%
```

```
DB: 3 Gaps: 2
US-10-089-688-1 (1-1053) x RAY94248 (1-341)
QY 1 ATCGCGTTCCCGTTTGGGAAGTCTCACAAATCTCCAGCAGACATTTGTGAAGAATCTGAAG 60
D 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTCTCGAAAGACAGACATTTCTGATAAAAAAGCAGAAAGCGCTTACA 120
D 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAAATCTGTTGCCATGAAGAAATCTGTATGGCAGCAAAATGAAAAA 180
D 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrglyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTCTCTCAAGAACTCTATATAGTGGGCTCCTT 240
D 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrglyThrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGTGTAGCTGATTACAGCTCATTTGACCTTTGAGGGCAAAAAAGACGTGGCTCAA 300
D 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACATATTTCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTTGAATACATC 360
D 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTygIle 120
QY 361 TGCAACCCAAAGAAATATTTTCTCATGTTATTTGAAAGGTATGAATCTCCAGAAATAGCT 420
D 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrglySerProGluIleAla 140
QY 421 CTAATTTGTGAATTAATTTAAGAAATGCAATCATCAGACATGAACCACTTGCAGAAATCAT 480
D 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTGGTCCGACAGATTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATA 540
D 161 LeuTrpSerGluGlnPheTyrglyAspPheArgTyrglyValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAAGACATAAATCTCAGTGCA 600
D 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTGGACAGCAATTTATGATGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
D 201 GluPheLeuGluGlnHisTyrglyAspArgPheSerGluTyrglyLysLeuLeuHisSer 220
QY 661 GAAATTTATGTGACAAAAGACAGTCACTGAGCTTCTCGGTGAACCTACTACTAGATAGA 720
D 221 GluAsnTyrglyThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAATTTCAATTTATGACAAAATACATCATGATAAATCTGAGAACCTCAATTAATGATG 780
D 241 HisAsnPheThrIleMetThrLysTyrglySerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACCTGTCTGCAGACAAAAGTCGCAACATCAGTTTTCAGGGCTTTTCAGGTTTTTAAGGTG 840
D 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCAAATCTTAACAGACGAGCCCAATCTTAGACATCTCTCTCAAGAACACAGGCC 900
D 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
QY 901 AAATCTATAGATTTCTTCAGAAATTTTCAGACACAGCAGGAGGATTTGATGAGCAGTTCC 960
D 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCGACGACGAATCTCCGGGTGATTTACCGGTTAAACCGCGACGCGTGGATCAGG 1020
D 316 AspGluGlnPheAsnAspGluLysThrTyrglyValLysGln-----IleArg 331
QY 1021 GATTGAAGACAGCAGCTCAGCAAGAAGCT 1050
```

```
DB 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341
RESULT 7
ABG23844
ID ABG23844 standard; protein; 354 AA.
XX AC ABG23844;
XX 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #23835.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX OS food supplement; medical imaging; diagnostic; genetic disorder.
XX PN Homo sapiens.
XX W0200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS88031.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PS biodiversity.
XX PS Claim 20; SEQ ID NO 54203; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 354 AA;
Alignment Scores:
Pred. No.: 5,57e-159 Length: 354
Score: 1597.50 Matches: 322
Percent Similarity: 92.57% Conservative: 2
Best Local Similarity: 92.00% Mismatches: 17
Query Match: 88.60% Indels: 9
DB: Gaps: 4
```

US-10-089-688-1 (1-1053) x ABG23844 (1-354)	
Qy	1 ATGCGGTTCCGTTTGGAGTCTCACAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
Db	14 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 33
Qy	61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAAGAGAGAAAGCGCTACA 120
Db	34 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 53
Qy	121 GAAGAAGTTTCCAAAAATCTGGTCCCATGAAGAAATTTCTGTATGGCACAATAAGAAAA 180
Db	54 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 73
Qy	181 GAGCTCAGACAGACAGTAGCTCAACTGCTCAAGAACTCTATATATAGTGGCTCTT 240
Db	74 AspProGlnThrGluAlaGlyAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 93
Qy	241 AGCACCTGGTAGCTGATTACAGCTCATTCAGCTTTCAGGGCAAAAGAGCTGGCTCAA 300
Db	94 IleThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 113
Qy	301 ATTTTCAACATAATTTCTCAGAAGCAAAATTCGTACGAGAACTCTACTGTTGAATACATC 360
Db	114 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 133
Qy	361 TGCACCCACAGATATTTCTTATGTTTATGTTTAAAGGGTATGATCTCCAGAAATAGCT 420
Db	134 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 153
Qy	421 CTAATTTGTGAATAATCTTAAAGAGATGTCATCAGACATGAACCACTTGCAGAAATCATT 480
Db	154 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuGlyLysIleIle 173
Qy	481 TTGTTGTCGGAACAGTTTTATGATTTCTCAGATATGTCGAAATGTCACAAATTTGACATA 540
Db	174 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 193
Qy	541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATTAATTTCTCAGTGCA 600
Db	194 AlaSerAspAlaPheAlaThrPheLysGlyLeuLeuThrArgHisLysLeuLeuSerAla 213
Qy	601 GAATTTTGGACACGATTTATGATAGATTTTTCAGTGAATATGAGAGTACTTACTTATCA 660
Db	214 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 233
Qy	661 GAAATATATGTCACAAAAGACAGTCACTAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db	234 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 253
Qy	721 CACAACTTCACATTTATCACAATAATACATCAGTAACTGAGAACCTCAAAATTTATGATG 780
Db	254 HisAsnPheThrIleMetThrLysTyrIleSerLysProValAsnLeuLeuMetMet 273
Qy	781 AACCTGCTGCAGACAAAGTCGCAACATTCAGATTTGAGGCTTTCACGTTTAAAGGTG 840
Db	274 AsnLeuLeuAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysAla 293
Qy	841 TTTGTAGCCAAATCTTAACAAGACGAGCCCAATCTCTAGACATCTCTCTCAAGAACGAGCC 900
Db	294 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 313
Qy	901 AAATCATAGAGTTCTCTCAGCAAGTTTTCAGACGACAGCGAGGATTTGATGAGCATGCC 960
Db	314 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 328
Qy	961 GTACCGACGACGAATTCCTCGGGTTCGATTTACGCGTTTAAACCGGACGCGTGGATCAGG 1020
Db	329 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 344
Qy	1021 GATTTGAAGAGACAGCTCAGCAAGAGCT 1050
Db	345 AspLeuLysArgProAlaGlnGlnGluAla 354

RESULT 8

AA94247
ID AAY94247 standard; protein; 337 AA.

AC AAY94247;

XX 10-AUG-2000 (first entry)

XX Human calcium binding protein hCBP.

XX Human; calcium binding protein; cancer; inflammation; CBP;
KW reproductive disorder; autoimmune disorder; developmental disorder;
KW seizure disorder; immune disorder; infection.

XX Homo sapiens.

XX WC200029580-A1.

XX 25-MAY-2000.

XX 12-NOV-1999; 99WO-US027027.

XX 13-NOV-1998; 98US-00190965.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX WPI; 2000-387793/33.

XX N-PSDB; AAA27332.

XX Human hCBP protein, and the nucleic acid encoding it, useful for e.g.
diagnosis, prevention and treatment of cancers, immune, developmental or
reproductive disorders.

XX Claim 1; Fig 1; 72pp; English.

XX The present sequence is the human calcium binding protein hCBP. It was
obtained by screening a coronary artery smooth muscle cDNA library, from
which five overlapping nucleic acids were isolated, sequenced and
expressed to give the protein. The protein and the gene encoding it are
useful for the diagnosis and treatment of the following types of
disorder: cancers (such as adenocarcinomas), reproductive disorders (such
as infertility, ovulatory defects, endometriosis, disruptions of the
oestrus and menstrual cycles, polycystic ovary syndrome and ovarian
hyperstimulation), autoimmune disorders (such as benign prostatic
hyperplasia and prostatitis), developmental disorders (such as Cushing's
syndrome, muscular dystrophy and gonadal dysgenesis), hereditary
neuropathies, seizure disorders, immune disorders (such as AIDS,
allergies, anaemia, asthma, atherosclerosis, cholecystitis, Crohn's
disease, diabetes, Graves' disease, multiple sclerosis, psoriasis,
rheumatoid arthritis, scleroderma, Sjogren's syndrome and ulcerative
colitis), and viral, bacterial, fungal, parasitic, protozoal and
helminthic infections

XX Sequence 337 AA;

Alignment Scores:

Pred. No.:	2,28e-127	Length:	337
Score:	1297.50	Matches:	263
Percent Similarity:	85.26%	Conservative:	32
Best Local Similarity:	76.01%	Mismatches:	38
Query Match:	71.96%	Indels:	13
DB:	3	Gaps:	3

US-10-089-688-1 (1-1053) x AAY94247 (1-337)

Qy 1 ATGCGGTTCCGTTTGGAGTCTCACAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60

Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysLeuLys 22

Qy 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGTATATAAAGCAGAAAGCGCTACA 120

```

Db      23  AspAenLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy      121 GAAGAAGTTTCCAAATCTGGTCCATGAAAGAAATCTGTATGGCACAAATGAAAAA 180
Db      40  GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
Qy      181 GAGCTCAGACAGACAGTACCTCACTCTCAAGAACTCTATATAGTGGCTCCTT 240
Db      60  GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu 79
Qy      241 AGCACCTGGTAGCTGATTACAGCTCATTGACCTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db      80  ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
Qy      301 ATTTTCAACATATCTCAGAAAGCAAAATTCGGTACGAGAACTCCTACTCTGTTGAATACATC 360
Db      100 IlePheAenAenIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
Qy      361 TGCACCCCAACAGATATTTTCTGATGTTATGATGATGATGATGATGATGATGATGATGAT 420
Db      120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
Qy      421 CTAATTTGTGGAATAATGTTAAGAGATGTCATCAGACATGAACCACTTGCACAAATCATT 480
Db      140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
Qy      481 TTGTGGTCGGAACAGTTTTATGATTTCTCAGATATGTCGAATATGTCGAATATGTCGAATAT 540
Db      160 LeuPheSerAenGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
Qy      541 GCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATTAATTCCTCAGTGCA 600
Db      180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
Qy      601 GAATTTTGGACACAGCATTTATGATAGATTTTTCAGTGAATATGAGAGATTTACTTCATTCA 660
Db      200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
Qy      661 GAAATATATGTGACAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db      220 GluAenTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
Qy      721 CACAACCTCACAATATGACAAATATACATCAGTAAACCTGAGAACCTCAATTAATGATG 780
Db      240 HisAenPheAlaIleMetThrLysTyrIleSerLysProGluAenLeuLysLeuMetMet 259
Qy      781 AACCTGCTGCAGACAAAGTCGCAACATCCAGTTTGAGGCTTTTCAGCTTTTAAAGTG 840
Db      260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
Qy      841 TTTGTAGCCAACTCTAAACAGACGACGCCATCCCTAGACATCCCTCCTCAAGAACCCAGGCC 900
Db      280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAenGlnPro 299
Qy      901 AAATCATAGAGTTCCTCAGCAAGTTTCAGAACGACGACGAGGATTTGATGAGCAGTTCC 960
Db      300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
Qy      961 GTACCGACGAGATTCCTGGGTGATTTACGCGTTAAACCGCGGACGCGTGGGATCAGG 1020
Db      315 -----AspGluGlnPheAlaAspGluLysAenTyrLeuIleLysGlnIleArg 330
Qy      1021 GATTTGAAGACACAGCT 1038
Db      331 AspLeuLysLysThrAla 336

```

RESULT 9
 AAB82090
 ID AAB82090 standard; protein; 337 AA.
 XX AC
 AC AAB82090;
 XX

```

DT      26-JUN-2001 (first entry)
XX      Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP.
DE      Human; cerebroprotective; neuroprotective; vaccine;
XX      Gene Therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW      stroke; acute head trauma; multiple sclerosis; spinal cord injury.
XX      Homo sapiens.
OS      WO200123552-A1.
XX      05-APR-2001.
XX      18-SEP-2000; 2000WO-EP009132.
XX      24-SEP-1999; 99EP-00118848.
XX      (MERE ) MERCK PATENT GMBH.
XX      Den Daas I, Duecker K;
XX      WPI; 2001-308142/32.
XX      N-PSDB; AAF86462.
XX      Novel human acute neuronal induced calcium binding polypeptide, and
PT      polynucleotides encoding them useful for diagnosing or treating stroke,
PT      acute head trauma, multiple sclerosis and spinal cord injury.
XX      Claim 1; Page 41-42; 45pp; English.
XX      The present sequence is the protein sequence for human Acute Neuronal
CC      Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
CC      protein are useful for treating stroke, acute head trauma, multiple
CC      sclerosis and spinal cord injury. ANIC-BP coding sequence and protein are
CC      also useful as vaccines for inducing an immunological response in a
CC      mammal
XX      SQ Sequence 337 AA;

```

Alignment Scores:
 Pred. No.: 2.28e-127 Length: 337
 Score: 1297.50 Matches: 263
 Percent Similarity: 85.26% Conservative: 32
 Best Local Similarity: 76.01% Mismatches: 38
 Query Match: 71.96% Indels: 13
 DB: Gaps: 3

US-10-089-688-1 (1-1053) x AAB82090 (1-337)

```

Qy      1  ATCCGTTCCCGTTTGGGAAGTCTCAAAATCTCCAGACACATTTGTGAAGAAATCTGAAG 60
Db      4  MetProLeu--PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
Qy      61  GAGAGCATGCTGTTCTGGAAAAGACAGACATTTCTGATATAAAAGACAGAAAAGGCTACA 120
Db      23  AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy      121 GAAGAAGTTTCCAAATCTGGTCCATGAAAGAAATCTGTATGGCACAAATGAAAAA 180
Db      40  GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
Qy      181 GAGCTCAGACAGACAGTACCTCACTCTCAAGAACTCTATATAGTGGCTCCTT 240
Db      60  GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu 79
Qy      241 AGCACCTGGTAGCTGATTACAGCTCATTGACCTTTGAGGGCAAAAAGAGCTGGCTCAA 300
Db      80  ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
Qy      301 ATTTTCAACATATCTCAGAAAGCAAAATTCGGTACGAGAACTCCTACTCTGTTGAATACATC 360
Db      100 IlePheAenAenIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119

```

```
QY 361 TGCACCCACAGAAATATTGTTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY 421 CTAATTTGTGAATAATGTTAAGAGAATGCAATGACATGAACCACTTGCAGAAATCATT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY 481 TTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAATGTCGAATTTGACATTA 540
Db 160 LeuPheSerAanGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATGATTTGCCACATTCAGGATTACTTCAACACACATAAATCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGACAGCATATGATAGATTTTTCAGTGAATATGAGAAATGTTACTTCATCA 660
Db 200 AspPheLeuGluGlnAenTyrAspThrIlePheGluAspTyrGluLysLeuGlnSer 219
QY 661 GAAATTTATGTCACAAAAGACAGTCACTCAAGCTTCTCGGTGAACCTACTACTAGATA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACTTCACAAATATGACAAAATACATCATGTAACCTGAGAACCTCAAAATTAATGAT 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACCTGCTGCAGACAAAAGTCGCAACATCCAGTTTGAAGGCTTTTCAGCTTTTAAGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCAATCTTAACAGACGAGCCCATCTAGACATCTCTCAAGAACCCAGGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 AAATCTATAGAGTTCTCTCAGCAAGTTTTCAGAACGACAGGACGGATTGTATGAGCAT 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGACGAATTCCTCCGGTTCGATTATACCGGTTAAACCGCGACCGCTGGATCAG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
QY 1021 GATTGAAGACACAGCT 1038
Db 331 AspLeuLysLysThrAla 336
RESULT 10
AAM39078
ID AAM39078 standard; protein; 337 AA.
XX AC
AC AAM39078;
XX AC
XX 22-OCT-2001 (first entry)
XX DE
XX Human polypeptide SEQ ID NO 2223.
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
```

```
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58234.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Example 4; SEQ ID NO 2223; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX SQ Sequence 337 AA;
Alignment Scores:
Pred. No.: 2,28e-127 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: Gaps: 3
US-10-089-688-1 (1-1053) x AAM39078 (1-337)
QY 1 ATGCCGTTCCCGTTGGGAAGTCTCACAATCTCCAGCAGCATTTGGAAGATCTGAAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
QY 61 GAGAGCATGGCTGTTCTGGAAGAACAGACATCTTCTGATAAAAGCAGAAAGAGCTACA 120
Db 23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
QY 121 GAAGAAGTTTCCAAAAATCTGTTGCCATGAAGAAATTTCTGTATGGCACAATGAAAAA 180
Db 40 GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
QY 181 GAGCCTCAGACAGAGCAGTAGCTCAACTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuLysTyrSerSerGlyLeuLeu 79
QY 241 AGCACCTGGTAGCTGATTACAGCTCATTGAGGGCAAAAAGACGCTGGCTCAA 300
Db 80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
```


Qy	301	ATTTTCAACAATATCTCAGAAAGACAAATTCGGTACGAGAACTCTTACTGTGTGAATACATC	360
Db	100	IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle	119
Qy	361	TGCACCCCAACAGAATATTTTCTTCATGTATTGAAAGGATGATGAATCTCCAGAATAAGCT	420
Db	120	SerAlaHisProHisIleLeuPheMetLeuLeuIysGlyTyrGluAlaProGlnIleAla	139
Qy	421	CTAAATTTGTGGAATAATGTTTAAAGAGAATGCAATGACATGAACCACTTTCGAAAAATCATT	480
Db	140	LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaIysIleIle	159
Qy	481	TTTGTGTCGGAACACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACAAATTTGACATA	540
Db	160	LeuPheSerAsnGlnPheArgAspPhePheIysTyrValGluLeuSerThrPheAspIle	179
Qy	541	GCTTCAGATGATTTGCCACATTCGAAGATTTACTTACAAGACACATAAATTCCTCAGTGCA	600
Db	180	AlaSerAspAlaPheAlaThrPheIysAspLeuLeuThrArgHisIysValLeuValAla	199
Qy	601	GAATTTTGTGAACAGCATTTATGATAGATTTTTTCAGTGAATGAGAAGTTACTTCTCATTTCA	660
Db	200	AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluIysLeuLeuGlnSer	219
Qy	661	GAATAATTTGTGACAAAAGACAGCTACTGAAGCTTCTCGGTGAAGTACTACTAGATAGA	720
Db	220	GluAsnTyrValThrIysArgGlnSerLeuIysLeuLeuGlyIleLeuLeuAspArg	239
Qy	721	CACAACTTCACAATTTATGACAAAATACATCAGTAAACCTGAGAACCTTCAAAATTAATCATG	780
Db	240	HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuIysLeuMetMet	259
Qy	781	AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTTCAGGCTTTCACGTTTTTAAAGTG	840
Db	260	AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal	279
Qy	841	TTTGTAGCCAATCTCAACAGACCCAGCCCACTCTAGACATCTCTCTCAAGAACCCAGGCC	900
Db	280	PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro	299
Qy	901	AAACTCATAGATTCCTTCAGCAAGTTTTCAGAACCAACAGAGCGGATTTGTATGACGAGTTC	960
Db	300	LysLeuIleGluPheLeuSerSerPheGlnIysGluArgThrAsp-----	314
Qy	961	GTACCCGACGCAATTCCTCCGGTTCGATTTACGCGTTAAACCGCGGACGCGTGGATCAGG	1020
Db	315	-----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg	330
Qy	1021	GATTTGAAGACACGAGCT	1038
Db	331	AspLeuLysLysThrAla	336

RESULT 11
AAB94139

ID AAB94139 standard; protein; 289 AA.

XX AAB94139:

XX
DT 26-JUN-2001 (first entry)

DE Human protein sequence SEO ID NO:14408.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
KW

XX
OS Homo sapiens.XX
PN
EP1074617-A2.XX
PD 07-FEB-2007

07-1EB-2001.
XX
28-III.-2000. 2000EB-00115125
PF

23 OCT-2000; 200001F=0011010120
29-JUL.-1999; 99JP-00248036

PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-0018776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
DR

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 14408; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03165 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 289 AA;

Alignment Scores:	3.23e-112	Length:	289
Pred. No.:	Score:	Matches:	231
	1153.50	Conservative:	34
Percent Similarity:	85.53%	Mismatches:	23
Best Local Similarity:	77.78%	Indels:	9
Query Match:	63.98%	Gaps:	1
DB:	4		

US-10-089-688-1 (1-1053) x AAB94139 (1-289)

QY	148	ATGAAAGAAATTCGTATGTCACAAATGAAAGAGGCTCAGACAGAAAGCAGTAGCTCAA	207
Db	1	MetLysGluIleLeuCysGlyThrAsnGluLysGluProThrGluAlaValAlaGln	20
QY	208	CTTGCTCAAACTCTATAATAGTGGGCTCTTAGCACCCCTGGTAGCTGATTACAGCTC	267
Db	21	LeuAlaGlnGluLeuTyrSerSerGlyLeuLeuValThrLeuIleAlaAspLeuGlnLeu	40
QY	268	ATTGACTTTTCAGGGCAAAAAGAGCTGGCTCAAAATTTTCAACAATATTTCTCAGAGACAA	327
Db	41	IleAspPheGluGlyIyLysAspValThrGlnIlePheAsnAsnIleLeuArgArgGln	60
QY	328	ATTGGTACGAGACTCCTACTGTGTGATACATCTGCACCCCAACAGATATTTTGTTCATG	387
Db	61	IleGlyThrArgSerProThrValGluTyrrIleSerAlaHisProHisIleLeuPheMet	80
QY	388	TTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGGAATATGTTTAAAGAA	447
Db	81	LeuLeuLysGlyItyrGluAlaProGlnIleAlaLeuArgCysGlyIleMetLeuArgGlu	100

448 TGATCAGATGACATGACCACTGTCAGAAATCATTTGTGGTCGGAACAGTTTATGATTC 507
 Db CysileargHisGluProLeuValLysleileLeuPheSerAsnGlnPheArgAspPhe 120
 508 TTCAGATATCTCGAATGTCAACATTTGACATAGCTTCAGATGCATTTGCCACATTCAG 567
 Db PheLysTyrValGluLeuSerThrPheAspIleAlaSerAspAlaPheAlaThrPheLys 140
 568 GATTACTTCAAGACATAAATCTCTCAGTCGCAAAATTTTGAACAGCAATATGATAGA 627
 Db AsnLeuLeuThrArgHisLysValLeuValAlaAspPheLeuGluGlnAsnTyrAspThr 160
 628 TTTTTCAGTAATGAGAGTACTTTCATTCAGAAATATGTCAGAAAACAGACATCA 687
 Db IlePheGluAspTyrGluLysLeuLeuGlnSerGluAsnTyrValThrLysArgGlnSer 180
 688 CTGAAGCTTCTCGTGAACTACTACTAGATAGACACAACTTCAAAATATGACAAATAC 747
 Db LeuLysLeuLeuGlyGluLeuIleLeuAspArgHisAsnPheAlaIleMetThrLysTyr 200
 748 ATCAGTAACTGAGAACCTCAATTAATGATGAACCTGCTGCGAGACAAAGTCGCAAC 807
 Db IleSerLysProGluAsnLeuLysLeuMetMetAsnLeuLeuArgAspLysSerProAsn 220
 808 ATCAGTTTGGGCTTCTCAGTTTAAAGTGTGTAGCCATCTCAACAGACGACGAG 867
 Db IleGlnPheGluAlaPheHisValPheLysValPheValAlaSerProHisLysThrGln 240
 868 CCCATCTAGACATCTCTCAAGAACAGCCGCAAACTCATAGAGTTCTCTCAGCAAGTTT 927
 Db ProlleValGluIleLeuLysAsnGlnProlLysLeuIleGluPheLeuSerSerPhe 260
 928 CAGAACAGCAGGACGATTTGATGACGAGTTCCGTACCGACGAGAAATCCCGGTCGAT 987
 Db GlnLysGluArgThrAsp-----AspGluGlnPheAla 271
 988 TTACCGTTAAACCGCGGACGCTGGGATCAGGATTTGAGAGACCGAGCT 1038
 Db AspGluLysAsnTyrLeuIleLysGlnIleArgAspLeuLysLysThrAla 288
 RESULT 12
 ID ABB60392 standard; protein; 339 AA.
 XX AC ABB60392;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 7968.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX FN WO2001/1042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US009231.
 XX PR 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PND, Myers EW;
 XX WPI; 2001-656860/75.
 XX DR N-PSDB; ABL04495.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

Genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
 Disclosure; SEQ ID NO 7968; 21pp + Sequence Listing; English.
 The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16178-ABL30511), expressed DNA sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fp.wipo.int/pub/published_pct_sequences
 Sequence 339 AA;

Alignment Scores: Pred. No.: 5,63e-109 Length: 339
 Score: 1123.00 Matches: 221
 Percent Similarity: 84.95% Conservative: 50
 Best Local Similarity: 69.28% Mismatches: 40
 Query Match: 62.29% Indels: 8
 DB: Gaps: 5

US-10-089-688-1 (1-1053) x ABB60392 (1-339)
 QY 1 ATCCGTTTCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTCTGAGAAATCTGAAG 60
 Db 1 MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLeuLys 19
 QY 61 GAGCAGATGCGTCTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTTACA 120
 Db 20 GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln 36
 QY 121 GAAGAAAGTTTCCAAAATCTGTTGCCATGAAAGAAATCTGTATGGCACAATAAGAAAA 180
 Db 37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuTyrGlySerSerAspAla 56
 QY 181 GAGCCTCAGACAGAA---GCAGTAGTCTCACTTGTCTCAGAACTCTATATAGTGGGCTC 237
 Db 57 GluProProAlaAspTyrValValAlaGlnLeuSerGlnGluLeuTyrAsnSerAsnLeu 76
 QY 238 CTTAGCACCTGCTGAGTGTATTTACAGCTCATTGACCTTTGAGGCGCAAAAAGAGCTGGCT 297
 Db 77 LeuLeuLeuLeuIleGlnAsnLeuHisArgIleAspPheGluGlyLysLysHisValAla 96
 QY 298 CAAATTTTCAACAATATTTCTCAGAACAAATTTGGTACGAGAACTCTCTACTGTTGAATAC 357
 Db 97 LeuIlePheAsnValLeuArgGlnIleGlyThrArgSerProThrValGluTyr 116
 QY 358 ATCTGACCCACACAGATATTTTGTTCATGTTTGAAGGGTATGATCT-----CCA 411
 Db 117 IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTyrGluAspAlaHisPro 136
 QY 412 GAAATAGCTCTAAATTTGTGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCA 471
 Db 137 GluIleAlaLeuAsnSerGlyThrMetLeuArgGluCysAlaArgTyrGluAlaLeuAla 156
 QY 472 AAAATCAATTTTGTGTCGGAACAGTATTTATGATTTCTTCAGATATGTCGAAATGTCACA 531
 Db 157 LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTyrValGluValSerThr 176
 QY 532 TTTGACATAGCTTCAGATGATTTTCCACATTCAGGATTTTACTTACAAGACATAAATG 591
 Db 177 PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu 196
 QY 592 CTCAGTGCAGAAATTTTGGAAACAGCATTTATGATGATTTTTCAGTGAA---TATGAGAG 648
 Db 197 LeuCysAlaGluPheLeuAspAlaAsnTyrAspLysPhePheSerGlnHisTyrGlnArg 216
 QY 649 TTACTTCTATTCAGAAAATTTATGTGACAAAAGACAGCTACTGAGGCTTCTCGGTGAACTA 708

Alignment Scores:					
Pred. No.:	2,41e-108	Length:	339		
Score:	1117.00	Matches:	220		
Percent Similarity:	84.95%	Conservative:	51		
Best Local Similarity:	68.97%	Mismatches:	40		
Query Match:	61.95%	Indels:	8		
DB:	3	Gaps:	5		
US-10-089-688-1 (1-1053) x AAY94249 (1-339)					
Qy	1	ATGCGCTTCGGTTTGGAGTCTCAAAATCTCCAGCAGACATTGTGAAGAANTCTGAAG	60		
Db	1	MetProLeu---PheGlySerSerglnYserProValGluLeuValLysSerLeuLys	19		
Qy	61	GAGAGCATGCTGTTCTTGAAAAACAAGACATTTCTGTATAAAAAGCAGAAAAGCGCTACA	120		
Db	20	GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln	36,		
Qy	121	GAAGAAGTTTCCAAAAATCTGGTGGCCATGAAGAATAATCTGTATGCCAATAATGA AAAA	180		
Db	37	GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuHisGlySerSerAspAla	56		
Qy	181	GAGCTCAGACAGAA----GCAGTAGCTCAACTTCTCAAGAACCTCTATATATAGTGGGCTC	237		
Db	57	GluProProAlaAspTryValValAlaGlnLeuSerGlnGluLeutyAsnSerAsnLeu	76		
Qy	238	CTTAGCACCTGGTAGCTGATTTACAGCTCATTCGACTTGTGAGGCGCAAAAAGACGTGGCT	297		
Db	77	LeuLeuLeuLeuIleGlnAsnLeuHisArgLLeaspPheGluGlyLysLysHisValAla	96		
Qy	298	CAAATTTTCAACAAATATCTCAGAAAGACAAATTTGGTACGAGAACTCTACTGTGTAATAC	357		
Db	97	LeuIlePheAsnAsnLeuLeuArgGlnIleGlyThrArgSerProThrValGluTyTr	116		
Qy	358	ATCTGCACCACACAGAAATATTTTGTTCATGTTATTTGAAGGGTATGAACTCT-----CCA	411		
Db	117	IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTrGluAspAlaHisPro	136		
Qy	412	GAATAGCTCTAAATTTGCGNAATCTTAAGAGAATGCATCAGACATGAACCACTTGCA	471		
Db	137	GluIleAlaLeuAsnSerGlyThrMetLeuArgGlnCysAlaArgTrGluAlaLeuAla	156		
Qy	472	AAATCATTTTGTGTCGGAACAGTTTTATGATTTCTTCAGATATGCGAAATGTCACAA	531		
Db	157	LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTrValGluValSerThr	176		
Qy	532	TTTGACATAGCTTCAGATGCAATTTGCCACATTAACGATTTCATTCAAGACATAAATTG	591		
Db	177	PheAspileAlaSerAspAlaPheSerThrPhelysGluLeuLeuThrArgHisLysLeu	196		
Qy	592	CTCAGTCCAGAAATTTTGGAAACAGCATATATGATAGATTTTTCAGTGAA---TATGAGNAG	648		
Db	197	LeuCysAlaGluPheLeuAspAlaAsnTrYAspLysPhePheSerGlnHisTrYcrlnArg	216		
Qy	649	TTACTTCATTACAGAAAATATGTGACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTA	708		
Db	217	LeuLeuAsnSerGluAsnTrYValThrArgGlnSerLeuLysLeuGlyGluLeu	236		
Qy	709	CTACTAGATAGACAACTTCACAAATTTATGACAAAATATACATCAGTAAACCTGAGAACCTC	768		
Db	237	LeuLeuAspArgHisAsnPheThrValIwetThrArgTrileSerGluProGluAsnLeu	256		
Qy	769	AAATTATATGATGAACCTGCTCGGAGACAAAAGTCGCAACATCCAGTTTGAGGCCCTTCAC	828		
Db	257	LysLeuMetMetAsnMetLeuLysGlnLysSerArgAsnIleGlnPheGluAlaPheHis	276		
Qy	829	GTTTTTAAGGTGTTTGTAGCAATTCCTTAAACAAGACGACGCCCATCTCTAGACATCTCTCTC	888		
Db	277	ValPheLysValPheValAlaAsnProAsnLysProllysProIleLeuAspileuleuLeu	296		
Qy	889	AAGAACAGGCCCAACTCATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGAGACCGAT	945		

RESULT 13

	Qy	121	GAGGAGGATTCCTCAAAATACGGTGTCCACAGAGGAAGAAGAACATCAGATATGACACCAATGCAGAAA	189
	Db	37	GLuAspValSerLysAenLeuValSerIleLysAenMetLeuHisGlySerSerAspAla	56
	Qy	181	GAGCGCTCAGACAGAAA---GCAGTAGCTCAACTTCCTCAAGAACCTCTATAATAGTGGGGCTC	237
	Db	57	GLuProProAlaAspTyrfyrValValAlaGlnLeuSerGlnGluLeutyfyrAasnSerAasnLeu	76
	Qy	238	CTTAGCACCCCTGGTAGCTGATTTACAGCTCATTTGACTCTTGAGGCGCAAAAAAGACGTGGCT	297
	Db	77	LeuLeuLeuLeuIleGlnAanLeuHisArgIleaspPheGluGlyLysLysHisValAla	96
	Qy	298	CAAATTTTCAACAATATTCTCAGAAGACAAATTTGGTACGAGAACTCTCTACTGTTGTAATPAC	357
	Db	97	LeuilePheAanAenLeuLeuArgGlnIleGlyThrArgSerProThrValGluIutyr	116
	Qy	358	ATCTGCACCACCAAGAAATATTTTGGTTCATGTTATTGAAAGGGTATGAATCT-----CCA	411
	Db	117	IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlytyrGluAspAlaHisPro	136
	Qy	412	GAATAGCTCTAAATTTGCGNAATATCTTAAGAGAATGCATCAGACATGAACCACTTGCA	471
	Db	137	GLuIleAlaLeuAenSerGlyThrMetLeuArgGluCysAlaArfytyrGluAlaLeuAla	156
	Qy	472	AAAATCATTTTGGTGGCGGAACAGTTTTATGATTTCTTCAGATATGTCGAATATGTCACA	531
	Db	157	LysIleMetLeuHisSerAspGluPhePheLysPhePheargtyrValGluValserThr	176
	Qy	532	TTTGACATAGCTTCAGATGCAATTTGCCACATTTCAAGGATTTTACTTCAAGACATATAATTG	591
	Db	177	PheAspilealaserAspAlaPheSerThrPhelysGluLeuleuthrArgHisLysLeu	196
	Qy	592	CTCAGTCAGAAATTTTGGAACACGATTTATGATAGATTTTTTCAGTGAA---TATGAGAGAG	648
	Db	197	LeuCysAlaGluPheLeuAenAspAlaAasnTyfyrAsplysPhePheSerGlnHisTyfyrGlnArg	216
	Qy	649	TTACTTCATTCAGAAAATATGTGACAAAAAGACAGCAGTCACGTGAAGCTTCTCGGTGAACCTA	708
	Db	217	LeuLeuAenSerGluAenTyfyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeu	236
	Qy	709	CTACTAGATAGACACACTTTCACAAATTTATGACAAAATACATCATCAGTAAACCTTGAGAACCTC	768
	Db	237	LeuLeuAspArgHisAenPheThrValMetThrArgtyrIleSerGluProGluAenLeu	256
	Qy	769	AAATTAATGATGAACCTGCTCGGAGACAAAAGTCGCAACATCCAGTTTGGAGCCCTTTTCAC	828
	Db	257	LysLeuMetMetAenMetLeuLysGlyLysSerArgAasnIleGlnPheGluAlaPheHis	276
	Qy	829	GTTTTAAAGTGTGTTAGCGCAATCCCTAAACAAAGACGCGCCCACTCTCCTAGACATCTCCTC	888
	Db	277	ValPheLysValPheValAlaAenProAasnLysBProLysProIleLeuAspIleLeuLeu	296
	Qy	889	AAGAACCGGCCCAAACTCATAGAGTTTCCTCAGCAAGTTTTCAGAACGACGAGCAGGAT	945

Db 297 ArgAsnGlnThrLysLeuValAspPheLeuThrAsnPheHisThrAspArgSerGlu 315

RESULT 14

AA94250

ID AAY94250 standard; protein; 377 AA.

XX

AC AAY94250;

XX

DT 10-AUG-2000 (first entry)

XX

DE C. elegans yeast-like calcium binding protein.

XX

KW Calcium binding protein; cancer; inflammation; yeast-like CBP; CBP;

KW reproductive disorder; autoimmune disorder; developmental disorder;

KW seizure disorder; immune disorder; infection.

XX

OS Caenorhabditis elegans.

XX

PN WO200029580-A1.

XX

PD 25-MAY-2000.

XX

PF 12-NOV-1999; 99WO-US027027.

XX

PR 13-NOV-1998; 98US-00190965.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX

DR WPI; 2000-387793/33.

XX

PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.

PT diagnosis, prevention and treatment of cancers, immune, developmental or

PT reproductive disorders.

XX

PS Disclosure; Page 68-69; 72pp; English.

XX

CC The present sequence is the C. elegans yeast-like CBP. It was used in a

CC sequence alignment to identify human calcium binding protein hCBP. The

CC hCBP protein and the gene encoding it are useful for the diagnosis and

CC treatment of the following types of disorder: cancers (such as

CC adenocarcinomas), reproductive disorders (such as infertility, ovulatory

CC defects, endometriosis, disruptions of the oestrus and menstrual cycles,

CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune

CC disorders (such as benign prostatic hyperplasia and prostatitis),

CC developmental disorders (such as Cushing's syndrome, muscular dystrophy

CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,

CC immune disorders (such as AIDS, allergies, anaemia, asthma,

CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'

CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis,

CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,

CC bacterial, fungal, parasitic, protozoal and helminthic infections

XX

SQ Sequence 377 AA;

Alignment Scores:

Pred. No.:	9, 71e-102	Length:	377
Score:	1054.50	Matches:	205
Percent Similarity:	77.44%	Conservative:	49
Best Local Similarity:	62.50%	Mismatches:	61
Query Match:	58.49%	Indels:	13
DB:	3	Gaps:	2

US-10-089-688-1 (1-1053) x AAY94250 (1-377)

QY 1 ATGCCGTTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATTGTGAAGAATCTGAAG 60

Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20

QY 61 GAGACATGGCTGTCTCGAAG-----CAAGAC 90

Db 21 GluValLeuThrIleLeuAspLysLeuProProProLysLeuAspLysAspGlyAsnIle 40

QY 91 ATTTCTGATAAAAAGCAGAAAAAGGCTACAGAAGTTCCTCAAAATCTGGTTGCCATG 150

Db 41 GlnSerAspLysLysTyAspLysAlaLeuAspGluValSerLysAsnValAlaMetIle 60

QY 151 AAAGAAATTTCTGTATGGCACAATGAAGAGAGCTTCAGACAGAA-----GCAGTA 201

Db 61 LysSerPheIleTyrgLysAsnAspSerAlaGluProSerGluHisValValGlnVal 80

QY 202 GCTCAACTTCTCAGAACTCTATAATAGTGGGCTCTTAGCACCCTGGTAGCTGATTTA 261

Db 81 AlaGlnLeuAlaGlnGluValTyAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100

QY 262 CAGCTCATCTTCTGAGGGCAAAAGACGTGCTCAAAATTTTCAACAATATTCTCAGA 321

Db 101 ProLysPheGluPheGluCysLysAspValGlyGlnIlePheAsnAsnLeuLeuArg 120

QY 322 AGACAAATTCGTACGAGAACTCTTACTGTTGAATACATCTGCACCCCAAGAAATTTTG 381

Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 140

QY 382 TTCATGTTATTGAAGGTATGAATCTCCAGAATAGCTCTAAATTTGTGTCGGAATAATCTTA 441

Db 141 IleGlnLeuValGlnGlyTySerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160

QY 442 AGAGAATGCATCAGACATGAACCACTTGCAAAATCATTTTGTGTCGGAACAGTTTAT 501

Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTySerAspValPheTy 180

QY 502 GATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGATTTGCCACA 561

Db 181 ThrPheLeuTyValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200

QY 562 TTCAGAGATTTACTTCAAGACATTAATTTCTCAGTGCAGAAATTTTGGACAGCATTAT 621

Db 201 PheLysGluLeuThrThrArgHisLysAlaIleIleAlaGluPheLeuAspSerAsnTy 220

QY 622 GATAGATTTTTCAGTGAATATGAGAAGTTACTTTCATTTCAGAAAATTTATGTGACAAAAGA 681

Db 221 AspThrPheAlaGlnTyrgLysAsnLeuLeuAsnSerLysAsnTyValThrArgArg 240

QY 682 CAGTCACTGAAGTTCTCGGTGAACTACTACTAGTAGACACAACTTCACAAATTTATGACA 741

Db 241 GlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArgHisAsnPheAsnThrMetThr 260

QY 742 AATATACATCAATAAACCCTGAGAACCTCAAAATTAATGATGAACCTGTCGCGAGACAAAGT 801

Db 261 LysTyrlleSerAsnProAspAsnLeuArgLeuMetGluLeuLeuArgAspLysSer 280

QY 802 CGCAACATCCAGTTTGAAGGCTTTTCAGTGTTTTAAAGTGTTTGTAGCCAAATCTCAACAAG 861

Db 281 ArgAsnIleGlnTyrgLysAlaPheHisValPheLysValPheValAlaAsnProAsnLys 300

QY 862 ACGAGGCCATCTCAGACATCTCTCAAGAACCGGCCAAACTCATAGATTTCTCAGC 921

Db 301 ProLysProIleSerAspIleLeuAsnArgAsnArgGluLysLeuValGluPheLeuSer 320

QY 922 AAGTTTTCAGAACGACGACGCGAT 945

Db 321 GluPheHisAsnAspArgThrAsp 328

RESULT 15

AA40864

ID AA40864 standard; protein; 237 AA.

XX

AC AA40864;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 5795.

XX

KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US034263.

23-DEC-1999; 99US-00471275.

21-JAN-2000; 2000US-00488725.

25-APR-2000; 2000US-00552317.

20-JUN-2000; 2000US-00598042.

19-JUL-2000; 2000US-00620312.

03-AUG-2000; 2000US-00653450.

14-SEP-2000; 2000US-00862191.

19-OCT-2000; 2000US-00693036.

29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Pi Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;

WPI; 2001-442253/47.

N-PSDB; AAI60020.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 2; SEQ ID NO 5795; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Sequence 237 AA;

Alignment Scores:	2.55e-84	Length:	237
Pred. No.:	888.50	Matches:	178
Score:	83.54%	Conservative:	20
Percent Similarity:	75.11%	Mismatches:	34
Best Local Similarity:	49.28%	Indels:	5
Query Match:	4	Gaps:	2

US-10-089-688-1 (1-1053) x AAM40864 (1-237)

Qy	334	ACGAGACTCTACTGTGTAATACATCTGCACCAACAGAAATATTTGTCATGTTATG 393
Db	2	ThrArgSerProThrValGluTyrIleSerAlaHisProHisIleLeuPheMetLeu 21
Qy	394	AAAGGGTATGATCTCCAGAAATAGCTCTAAATGTCGAATATGTTAAGAGATGCATC 453
Db	22	LysGlyTyrGluAlaProGlnIleAlaLeuArgCysGlyIleMetLeuArgGluCys 41

Qy	454	AGACATGAAACCACTTGCAAAAATCAATTTGTTGGTCGGAAACAGTTTATGATTTCTTCAGA 513
Db	42	ArgHisGluProLeuAlaLysIleIleLeuPheSerAenGlnPheArgAspPhePheLys 61
Qy	514	TATGTCGAAATGTCAACATTTGACATAGCTTTCAGATGATTTGCGACATTCACAGGATTTA 573
Db	62	TyrValGluLeuSerThrPheAspIleAlaSerAspAlaPheAlaThrPheLysAspLeu 81
Qy	574	CTTACAAAGACATAAATGTCTCAGTGCAGAAATTTTGGACAGACATATGATAGATTTC 633
Db	82	LeuThrArgHisValLeuValAlaAspPheLeuGluGlnAenTyrAspThrIlePhe 101
Qy	634	AGTGAATATGAGAAGTTACTTTCATTCAGAAAATTTATGTCAGAAAACAGACAGTCTAGAG 693
Db	102	GluAspTyrGluLysLeuGlnSerGluAenTyrValThrLysArgGlnSerLeuLys 121
Qy	694	CTTCTCGGTGAATCTACTACTAGATAGACACAACTTTCACAATTTATGACAAAATACATCAGT 753
Db	122	LeuLeuGlyGluLeuIleLeuAspArgHisAsnPheAlaIleMetThrLysTyrIleSer 141
Qy	754	AAACCTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAATCCAG 813
Db	142	LysProGluAenLeuLysLeuMetMetAsnLeuLeuArgAspLysSerProAsnIleGln 161
Qy	814	TTTGAGCCCTTTCACGTTTAAAGTGTGTTGAGCCAAATCCTAACAGAGCGAGCCCATC 873
Db	162	PheGluAlaPheHisValPheLysValPheValAlaSerProHisLysThrGlnProIle 181
Qy	874	CTAGACATCTCTCAAGAACCCAGGCCCAACTCATAGAGTTCTCTCAGCAAGTTTCAGAAC 933
Db	182	ValGluIleLeuLeuLysAenGlnProLysLeuIleGluPheLeuSerSerPheGlnLys 201
Qy	934	GACAGGACGATTTGATGAGCAGTTTCGTACCGACGACGAATTCCTCCGGGTC----- 984
Db	202	GluArgThrAsp---AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIle 220
Qy	985	---GATTACGCGTTAAACCCCGGACCGCGGATCAGGATCAGGATTTGAAGAGA 1032
Db	221	ArgAspLeuLysLysThrAlaPro***ArgAlaLeuArgAspSerLysArg 237

Search completed: November 10, 2004, 20:43:03
Job time : 180.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:34:48 ; Search time 36.5 Seconds
(without alignments)
3826.458 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atgcgcgtccctgttgga.....cagctcgaagaagcttaa 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Deiop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model
-Q=/cgn2_1/USPRO.spool_p/US10089688/runat_10112004_152247_4873/app_query.fasta_1.1223
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN_1_1_46 @runat_10112004_152247_4873 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/FACTUS COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1614.5	89.5	341	3	US-09-190-965-3
2	1614.5	89.5	341	3	US-09-470-253-3
3	1297.5	72.0	337	3	US-09-190-965-1
4	1297.5	72.0	337	3	US-09-470-253-1
5	1117	62.0	339	3	US-09-190-965-4
6	1117	62.0	339	3	US-09-470-253-4
7	1054.5	58.5	377	3	US-09-190-965-5
8	1054.5	58.5	377	3	US-09-470-253-5
9	358.5	19.9	165	4	US-09-248-796A-14303
10	166	9.2	125	4	US-09-248-796A-15202
11	163	9.0	90	4	US-09-248-796A-15201
12	115.5	6.4	630	4	US-09-248-796A-20275

13	113	6.3	541	4	US-09-134-000C-5420	Sequence 5420, Ap
14	113	6.3	2184	4	US-09-417-485D-6	Sequence 6, Appli
15	110.5	6.1	533	3	US-09-134-001C-4053	Sequence 4053, Ap
16	106	5.9	3959	3	US-08-970-269A-30	Sequence 30, Appl
17	106	5.9	3959	3	US-09-407-562-30	Sequence 30, Appl
18	105.5	5.9	3878	4	US-09-914-259-11	Sequence 11, Appl
19	105	5.8	478	4	US-09-328-352-6512	Sequence 6512, Ap
20	105	5.8	496	4	US-09-543-681A-6485	Sequence 6485, Ap
21	105	5.8	627	1	US-09-248-796A-20614	Sequence 20614, A
22	105	5.8	776	1	US-08-021-601-2	Sequence 2, Appli
23	105	5.8	776	5	PCT-US94-01624-2	Sequence 2, Appli
24	105	5.8	776	5	PCT-US94-01624-2	Sequence 2, Appli
25	104	5.8	933	4	US-09-248-796A-20513	Sequence 20513, A
26	103.5	5.7	1388	3	US-09-572-191-2	Sequence 2, Appli
27	103.5	5.7	1388	3	US-09-723-262-2	Sequence 2, Appli
28	103.5	5.7	1388	3	US-09-723-219-2	Sequence 3, Appli
29	103.5	5.7	2474	4	US-08-265-967C-3	Sequence 3, Appli
30	103.5	5.7	2474	4	US-08-305-790B-4	Sequence 4, Appli
31	103	5.7	1155	4	US-09-543-681A-6286	Sequence 6286, Ap
32	103	5.7	1388	2	US-08-685-576-4	Sequence 4, Appli
33	102.5	5.7	670	3	US-09-134-001C-2940	Sequence 2940, Ap
34	102	5.7	506	4	US-09-107-532A-5994	Sequence 5994, Ap
35	101.5	5.6	467	4	US-09-543-681A-4262	Sequence 4262, Ap
36	101	5.6	489	4	US-09-248-796A-19030	Sequence 19030, A
37	100.5	5.6	1388	2	US-08-685-576-1	Sequence 1, Appli
38	100	5.5	460	3	US-08-942-012B-33	Sequence 33, Appl
39	100	5.5	831	4	US-09-269-861A-8	Sequence 8, Appli
40	99	5.5	2954	4	US-09-150-867-1	Sequence 1, Appli
41	98.5	5.5	568	4	US-09-710-279-2482	Sequence 2482, Ap
42	98.5	5.5	967	4	US-09-543-681A-6407	Sequence 6407, Ap
43	98.5	5.5	1145	4	US-09-710-279-1432	Sequence 1432, Ap
44	98.5	5.5	1151	3	US-09-134-001C-3242	Sequence 3242, Ap
45	98	5.4	1388	4	US-09-976-594-296	Sequence 296, App

ALIGNMENTS

RESULT 1
US-09-190-965-3
; Sequence 3, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PP-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-190-965-3

Alignment Scores:
Pred. No.: 8.34e-173 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-190-965-3 (1-341)

QY 1 ATGCGTTCCTCCGTTGGAGTCTCACAAATCTCCACACAGCATTTGTGAGAATCTGAAG 60
|||||
Db 1 MetcProPheProGlyLysSerHisLysSerProAlaHisValLysAsnLeuLys 20

```

120 QY 61 GAGAGCATGGCTGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTTACA
121 Db 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr
122 QY 121 GAGAAGTTTCCAAAATCTGGTCCCATGAAAGAAATCTGTATGGCACAATGAAAAA
123 Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLeuTyrGlyThrAsnGluLys
124 QY 181 GAGCCTCAGACAGAGTAGCTCAACTTCTCAAGAACTCTATATATAGTGGCTCCTT
125 Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu
126 QY 241 AGCACCTGGTAGCTGATTACAGCTCATGACTTTCAGGGCAGAAAAGAGCTGGCTCAA
127 Db 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln
128 QY 301 ATTTTCAACAAATATCTCAGAGCAAAATGGTACGAGAACTCTCTACTGTTGAATACATC
129 Db 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle
130 QY 361 TGCACCCCAACAAGATATTTTTCATGTTAATGAAAGGTTATGAATCTCCAGAAATAGCT
131 Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla
132 QY 421 CTAAATTTGGAATTAATGTTAAGAGATGCAATGACATGACACCACTTGCAGAAATCATT
133 Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle
134 QY 481 TTGTGGTCCGAAACAGTTTATGATTCTTCAGATATGTCCAAATGTCAACATTTGACATA
135 Db 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle
136 QY 541 GCTTCAGATGCATTTGCCACATTCAGGATTTAATTACAGACATTAATCTCAGTGCA
137 Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuSerAla
138 QY 601 GAAATTTTGGACAGCATATGATAGATTTTTCAGTGAATATGAGAGTTACTTCAATCA
139 Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer
140 QY 661 GAAATTTATGTGACAAAAGACAGTCACTCAAGCTTCTCGGTGAACCTACTACTAGATAGA
141 Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg
142 QY 721 CACAACTTCAATATGACAAATATACATACATGTAACCTGAGAACCTCAAATTAATGATG
143 Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet
144 QY 781 AACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGAAGGCTTTCACGTTTTTAAGGTG
145 Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal
146 QY 841 TTTGTAGCCAAATCCTAACAAGACGAGCCCATCTCTAGACATCTCTCTCAAGAACCCAGGCC
147 Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr
148 QY 901 AAATCTATAGATTTCTCAGCAAGTTTCAGAAAGCAGACGCGATGTTGATGACGATGCC
149 Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu
150 QY 961 GTACCGACGCAATTCGCGGTTCGATTACGCGTTAAACCGCGACGCGTGGGATCAGG
151 Db 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
152 QY 1021 GATTTGAGAGACAGCTCAGCAAGAAGCT 1050
153 Db 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341
```

RESULT 2
US-09-470-253-3
; Sequence 3, Application US/09470253
; Patent No. 6365371

```

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: g262934
US-09-470-253-3

Alignment Scores:
Pred. No.: 8,34e-173 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-470-253-3 (1-341)
QY 1 ATCCGCTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGTGAAGAATCTGAAG 60
Db 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTTACA 120
Db 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAGAGATTTCCAAAATCTGTTGCCATGAAAGAAATCTGTATGGCACAACAAATGAAAAA 180
Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGAGTAGCTCAACTTCTCAAGAACTCTATATATAGTGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTAGCTGATTACAGCTCATGACTTTCAGGGCAGAAAAGAGCTGGCTCAA 300
Db 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACAAATATCTCAGAGCAAAATGGTACGAGAACTCTCTACTGTTGAATACATC 360
Db 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCCAACAAGATATTTTTCATGTTAATGAAAGGTTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGGAATTAATGTTAAGAGATGCAATGACATGACACCACTTGCAGAAATCATT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTGGTCCGAAACAGTTTATGATTCTTCAGATATGTCCAAATGTCAACATTTGACATA 540
Db 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCATTTGCCACATTCAGGATTTAATTACAGACATTAATCTCAGTGCA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuSerAla 200
QY 601 GAAATTTTGGACAGCATATGATAGATTTTTCAGTGAATATGAGAGTTACTTCAATCA 660
Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer
```

```
Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAATATATGTGACAAAAGACAGTCACTAGAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyGluLeuLeuLeuAspArg 240
Qy 721 CACAACTTCACAATTTATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATCATG 780
Db 241 HisAsnPhenThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTTCAGGCGCTTTCACGTTTTTAAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCAATCTTACACAGACGAGCCATCTCTAGACATCTCTCTCAAGAACACGAGCC 900
Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAACCTCATAGAGTTCCCTCAGCAAGCTTTCAGAACACGAGGAGCGGATTGTATGAGCAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 331
Qy 961 GTACCGACGACGAATTCCTCGGTCGATTTCAGCGCTTAAACCGCGAGCGGTGGATCAGG 1020
Db 331 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331

RESULT 3
US-09-190-965-1
; Sequence 1, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-190-965-1

Alignment Scores:
Pred. No.: 3,89e-137 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: 3 Gaps: 3

US-10-089-688-1 (1-1053) x US-09-190-965-1 (1-337)
Qy 1 ATGCGGTCCCGTTTGGAGTCTCACAAATCTCCAGCAGACATTTGGAAGATCTGAAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
Qy 61 GAGAGCATGCTGTCTCGAAAAGCAAGACATTTCTCATAAAAAGCAGAAAAGGCTACA 120
Db 23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy 121 GAAGAAATTTCCAAAATCTGGTCCCATGAAGAAATTTCTGTATGCGCACAAATGAAAA 180
```

```
Db 40 GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATACTAGTGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuTyrSerSerGlyLeuLeu 79
Qy 241 AGCACCTCGTAGCTGATTTACAGCTCATTGACCTTTGAGGGCAAAAAGACGCTGCTCAA 300
Db 80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
Qy 301 ATTTTCAACCAATATTTCTCAGAACAAATTTGGTACGAGAACTCTCTACTGTTGAATACATC 360
Db 100 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
Qy 361 TGCACCCCAACAGAAATATTTGTTTCATGTTTAAAGGGTATGAATCTCCAGAAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLysGlyTyrGluAlaProGlnIleAla 139
Qy 421 CTAAATTTGTAATATGTTAAAGAAATGCAATCAGACATCAACCACTTCGAAAATCATT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
Qy 481 TTGTGTCGGAACAGATTTTATGATTTTCTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 160 LeuPheSerAsnGlnPheArgAspPheLysTyrValGluLeuSerThrPheAspIle 179
Qy 541 GCTTCAGATGATTTGCCACATTTCAAGGATTTTACTTACAGACATAAATTTGCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
Qy 601 GAATTTTGGACACAGCATTTATGATATGATTTTTCAGTGAATATGAGAAGTTACTTCATCA 660
Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuGlnSer 219
Qy 661 GAAAAATTATGTGACAAAAGACAGTCACTGCAAGCTTCTCGTCAACTTACTACTAGATAGA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
Qy 721 CACAACTTCACAATTTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATG 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
Qy 781 AACCTGCTGCGAGACAAAAGTCCCAATCCAGTTTTCAGGCTTTCAGCTTTTAAAGGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
Qy 841 TTTGTAGCCCAATCTTAAACAGACGAGCCCATCTAGACATCTCTCTCAAGAACACGAGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
Qy 901 AAACCTCATAGAGTTCTCAGCAAGTTTTCAGAACGAGCAGGACGAGGATTTGTATGAGCAGTTCC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
Qy 961 GTACCGACGAGCAATTCCTCGGTCGATTTACGGGTAAACCGCGAGCGGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrLeuIleLysGlnIleArg 330
Qy 1021 GATTTCAAGACAGACAGCT 1038
Db 331 AspLeuLysLysThrAla 336

RESULT 4
US-09-470-253-1
; Sequence 1, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
```

/ CURRENT APPLICATION NUMBER: US/09/470,253	
/ CURRENT FILING DATE: 1999-12-22	
/ PRIOR APPLICATION NUMBER: 09/190,965	
/ PRIOR FILING DATE: 1998-11-13	
/ NUMBER OF SEQ ID NOS: 5	
/ SOFTWARE: PERL Program	
/ SEQ ID NO 1	
/ LENGTH: 337	
/ TYPE: PRT	
/ ORGANISM: Homo sapiens	
/ FEATURE: -	
/ OTHER INFORMATION: 3734805	
US-09-470-253-1	
Alignment Scores:	
Pred. No.:	3,89e-137
Score:	1297.50
Percent Similarity:	85.26%
Best Local Similarity:	76.01%
Query Match:	71.96%
DB:	3
US-10-089-688-1 (1-1053) x US-09-470-253-1 (1-337)	
Qy	1 ATGCGTCCCGTTGGGAAGTCTCAAAATCTCCAGACACATTTGTGAAGATCTGAAG 60
Db	4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluLeuValLysIleLeuLys 22
Qy	61 GAGAGCATGGCTGTCTCGAAAGACAGACATTTCTGTATAAAAGCAGAAAGCGGTACA 120
Db	23 AsnAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
Qy	121 GAAGAGTTTCCAAAATCTGGTTCGCAATGAAGAAATCTGTATGGCACAATAGAAAA 180
Db	40 GluGluValSerLysSerLeuGlnAlaValAlaGlnLeuAlaGlnLeuLysGlyThrAsnGluLys 59
Qy	181 GAGCTCAGACAGACAGTACCTCAACTGCTCAAGACTCTATAATAGTGGCTCCTT 240
Db	60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuLysSerSerGlyLeuLeu 79
Qy	241 AGCACCTGGTGTGATTACAGCTCATTGACCTTTGAGGGCAAAAGACGTGGCTCAA 300
Db	80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValThrGln 99
Qy	301 ATTTCAACATATCTCAGAAACAATGGTACGAGAACTCTACTGTGAATACATC 360
Db	100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluLysIle 119
Qy	361 TGCACCCCAACAGATATTTTCTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCT 420
Db	120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyThrGluAlaProGlnIleAla 139
Qy	421 CTAATTTGTGAATAATGTTAAGAAATGTCATCAGACATGAACCACTTGCAAAATCATT 480
Db	140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 159
Qy	481 TTGCGTCGGAACAGATTTTATGATTTCTTCATGATGTCGAATGTCACATTTGCACATA 540
Db	160 LeuPheSerAsnGlnPheArgAspPhePheLysThrValGluLeuSerThrPheAspIle 179
Qy	541 GCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAAGACATAAATTCGTCAATGCA 600
Db	180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
Qy	601 GAAATTTTGGACAGCATATGATAGATTTTTCAGTGAATATGAGAATTTACTTCAATCA 660
Db	200 AspPheLeuGluGlnAsnThrAspThrIlePheGluAspThrGluLysLeuGlnSer 219
Qy	661 GAAATTTATGCAAAAAGACAGTCACTGAAGCTTCTCGGTGGAAGTACTACTAGATAGA 720
Db	220 GluAsnThrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
Qy	721 CACAACCTTCAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
RESULT 5	
US-09-190-965-4	
/ Sequence 4, Application US/09190965	
/ Patent No. 6071721	
/ GENERAL INFORMATION:	
/ APPLICANT: Tang, Y. Tom	
/ APPLICANT: Guegler, Karl J.	
/ APPLICANT: Corley, Neil C.	
/ APPLICANT: Gorgone, Gina A.	
/ TITLE OF INVENTION: CALCIUM BINDING PROTEIN	
/ FILE REFERENCE: PF-0635 US	
/ CURRENT APPLICATION NUMBER: US/09/190,965	
/ CURRENT FILING DATE: 1998-11-13	
/ NUMBER OF SEQ ID NOS: 5	
/ SOFTWARE: PERL Program	
/ SEQ ID NO 4	
/ LENGTH: 339	
/ TYPE: PRT	
/ ORGANISM: Drosophila melanogaster	
/ FEATURE: -	
/ OTHER INFORMATION: g1794137	
US-09-190-965-4	
Alignment Scores:	
Pred. No.:	7.99e-117
Score:	1117.00
Percent Similarity:	84.95%
Best Local Similarity:	68.97%
Query Match:	61.95%
DB:	3
US-10-089-688-1 (1-1053) x US-09-190-965-4 (1-339)	
Qy	1 ATGCGTTCCTCCGTTTGGGAAGTCTCAAAATCTCCAGACACATTTGTGAAGATCTGAAG 60
Db	1 MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLeuLys 19
Qy	61 GAGAGCATGGCTGTCTCGAAAGACAGACATTTCTGTATAAAAGCAGAAAGCGGTACA 120
Db	20 GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln 36
Qy	121 GAAGAGTTTCCAAAATCTGGTTCGCAATGAAGAAATCTGTATGGCACAATAGAAAA 180
Db	37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuHisGlySerSerAspAla 56
Qy	181 GAGCTCAGACAGAA---GCAGTAGCTCAACTTCTCTCAAGAACTCTATATAGTGGGCTC 237
Db	57 GluProProAlaAspThrValValAlaGlnLeuSerGlnGluLeuTyAsnSerAsnLeu 76
Qy	238 CTTAGACCCCTGGTAGCTGATTACAGCTCATTGAGCTTTGAGGGCAAAAGACGTGGCT 297

```
Db 77 LeuLeuLeuLeuIleclnAsnLeuHi sArgIleAspPheGluGlyLysHisValala 96
Qy 298 CAAATTTTCAAAATTTCTCAGAGACAATAATGGTAGCAGAACTCTCTACTGTTGAATAC 357
Db 97 LeuIlePheAsnAsnLeuLeuArgGlnIleGlyThrArgSerProThrValGluTyr 116
Qy 358 ATCTGCACCAACAGCAATATTTGTTGATGTTATTGAAAGGGTATGAATCT-----CCA 411
Db 117 IleCysThrLysProGluLeuLeuPheThrLeuMetAlaGlyTyrGluAspAlaHisPro 136
Qy 412 GAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGCATCAGACATCAACCACTTGCA 471
Db 137 GluIleAlaLeuAsnSerGlyThrMetLeuArgGluCysAlaArgTyrGluAlaLeuAla 156
Qy 472 AAAATCATTTTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAATATGTCACACA 531
Db 157 LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTyrValGluValSerThr 176
Qy 532 TTTGACATAGCTTCAGATGCAATTCGCCACATTCAGAGATTTTACTTACAGACATAAATG 591
Db 177 PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu 196
Qy 592 CTCAGTCAGAAATTTTGGACACAGCATATGATAGATTTTTCAGTGAA---TATGAGAAG 648
Db 197 LeuCysAlaGluPheLeuAspAlaAsnTyrAspLysPhePheSerGlnHisTyrGlnArg 216
Qy 649 TTACTTCAATTCAGAAAATATGTCAGACAAATAATGGTAGCAGAACTCTCTCGGTGAACCTA 708
Db 217 LeuLeuAsnSerGluAsnTyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeu 236
Qy 709 CTACTAGATAGACACAACTTCACATATGATGACAAAATACATCAGTAAACCTGAGAACCTC 768
Db 237 LeuLeuAspArgHisAsnPheThrValMetThrArgTyrIleSerGluProGluAsnLeu 256
Qy 769 AAAATTAATGATGAACCTCTCGAGACAAAAGTCGCAACATCCAGTTTGGAGCCTTTTCAC 828
Db 257 LysLeuMetMetAsnMetLeuLysGluLysSerArgAsnIleGlnPheGluAlaPheHis 276
Qy 829 GTTTTAAAGTGTTGTTAGCCAATCTTAACAGACGCGCCCATCTAGACATCTCTCCTC 888
Db 277 ValPheLysValPheValAlaAsnProAsnLysProLysProIleLeuAspIleLeuLeu 296
Qy 889 AAGAACCGGCCCAACTCATAGATTTCTCAGCAAGTTTCTCAGACGACGACGACGAT 945
Db 297 ArgAsnGlnThrLysLeuValAspPheLeuThrAsnPheHisThrAspArgSerGlu 315
```

RESULT 6

```
US-09-470-253-4
; Sequence 4, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: gi794137
US-09-470-253-4
```

Alignment Scores:

```
Pred. No.: 7,99e-117 Length: 339
Score: 1117.00 Matches: 220
Percent Similarity: 84.95% Conservative: 51
Best Local Similarity: 68.97% Mismatches: 40
Query Match: 61.95% Indels: 8
DB: 3 Gaps: 5

US-10-089-688-1 (1-1053) x US-09-470-253-4 (1-339)
Qy 1 ATGCGTTTCCGGTTTGGGAAGTCTCACAATAATCTCCAGCAGACATTTGTGAAGAATCTGAAG 60
Db 1 MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLys 19
Qy 61 GAGAGCATGGCTGTTCTGGAAGAACAGACATTTCTGATATAAAAGCAGAAAAGGCTACA 120
Db 20 GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln 36
Qy 121 GAAGAAGTTTCCAAAATATGTTGCCATGAAGAAATCTGTATGGCACAATAATGAAAAA 180
Db 37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuHisGlySerSerAspAla 56
Qy 181 GAGCCTCAGACAGAA---GCAGTAGCTCAACTCTCAAGAACTCTATAATAGTGGGCTC 237
Db 57 GluProProAlaAspTyrValValAlaGlnLeuSerGlnGluLeuTyrAsnSerAsnLeu 76
Qy 238 CTTAGCACCTTGGTAGCTGATTTACAGCTCATTCAGCTTTGAGGCCAAAAGACGCTGGCT 297
Db 77 LeuLeuLeuLeuIleGlnAsnLeuHisArgIleAspPheGluGlyLysHisValAla 96
Qy 298 CAAATTTTCAACATATTTCTCAGNAGACAAATTTGGTAGGAACTCTCTACTGTTGAATAC 357
Db 97 LeuIlePheAsnAsnLeuLeuArgGlnIleGlyThrArgSerProThrValGluTyr 116
Qy 358 ATCTGCACCAACAGAAATATTTTTCATGTTTGAAGGGTATGAATCT-----CCA 411
Db 117 IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTyrGluAspAlaHisPro 136
Qy 412 GAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTTGCA 471
Db 137 GluIleAlaLeuAsnSerGlyThrMetLeuArgGluCysAlaArgTyrGluAlaLeuAla 156
Qy 472 AAAATCATTTTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAATATGTCACACA 531
Db 157 LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTyrValGluValSerThr 176
Qy 532 TTTGACATAGCTTCAGATGCAATTCGCCACATTCAGGATTTTACTTACAAGACATAAATG 591
Db 177 PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu 196
Qy 592 CTCAGTCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAA---TATGAGAAG 648
Db 197 LeuCysAlaGluPheLeuAspAlaAsnTyrAspLysPhePheSerGlnHisTyrGlnArg 216
Qy 649 TTACTTCAATTCAGAAAATATGTCAGACAAATAATGGTAGCAGAACTCTCTCGGTGAACCTA 708
Db 217 LeuLeuAsnSerGluAsnTyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeu 236
Qy 709 CTACTAGATAGACACAACTTCACATATGACAAAATACATCAGTAAACCTGAGAACCTC 768
Db 237 LeuLeuAspArgHisAsnPheThrValMetThrArgTyrIleSerGluProGluAsnLeu 256
Qy 769 AAAATTAATGATGAACCTCTCGAGACAAAAGTCGCAACATCCAGTTTGGAGCCTTTTCAC 828
Db 257 LysLeuMetMetAsnMetLeuLysGluLysSerArgAsnIleGlnPheGluAlaPheHis 276
Qy 829 GTTTTAAAGTGTTTGTAGCCAATCTTAACAGACGCGCCCATCTAGACATCTCTCCTC 888
Db 277 ValPheLysValPheValAlaAsnProAsnLysProLysProIleLeuAspIleLeuLeu 296
Qy 889 AAGAACCGGCCCAACTCATAGATTTCTCAGCAAGTTTCTCAGACGACGACGACGAT 945
Db 297 ArgAsnGlnThrLysLeuValAspPheLeuThrAsnPheHisThrAspArgSerGlu 315
```

```
RESULT 7
US-09-190-965-5
; Sequence 5, Application US/09190965
; Patent No. 6071721
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/190,965
; CURRENT FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-190-965-5

Alignment Scores:
Pred. No.: 9.06e-110 Length: 377
Score: 1054.50 Matches: 205
Percent Similarity: 77.44% Conservativeness: 49
Best Local Similarity: 62.50% Mismatches: 61
Query Match: 58.49% Indels: 13
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-190-965-5 (1-377)

QY 1 ATGCCGTTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGGAAGATCTGAAG 60
DB 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20
QY 61 GAGACATGGCTGTCGAAAG-----CAAGAC 90
DB 21 GluValLeuThrIleLeuAspLysLeuProProLysLysLeuAspGlyAsnIle 40
QY 91 ATTTCTGATAAAAGCAGAAAGCTACAGAAAGTTTCAAAATCTCGTTGCCATG 150
DB 41 GlnSerAspLysLysTyAspLysAlaLeuAspGluValSerLysAsnValAlaMetIle 60
QY 151 AAGAAATTCGTATGGCACAATGAAAGAGCTTCAGACAGAA-----GCAGTA 201
DB 61 LysSerPheIleTyArgSerAlaGluProSerSerGluHisValValGlnVal 80
QY 202 GCTCAACTGCTCAGAACTCTATAATAGTGGCTCTTAGCACCCTGTTAGCTGATTTA 261
DB 81 AlaGlnLeuAlaGlnGluValTyArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 100
QY 262 CAGCTCATTGACTTTGAGCGCAAAAGACGTGGCTCAAAATTTCAACAATATTTCTCAGA 321
DB 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnLeuLeuArg 120
QY 322 AGACAAATGGTACGAGAACTCTACTGTTGAATACATCTGCACCACCAAGAAATTTTG 381
DB 121 ArgGlnIleGlyThrArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 140
QY 382 TTCATGTTATTAAGAGGTGAATCTCCAGAAATAGCTCTAAATTTGGGAATATGTTA 441
DB 141 IleGlnLeuValGlnGlyTySerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160
QY 442 AGAAGATGATCAGACATGAACCACTTGCAGAAATCATTTGTTGTCGGAACAGTTTAT 501
DB 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTySerAspValPheTy 180
QY 502 GATTCTTCAGATATGTCGAATGTCAACATTTTGACATAGCTTCAGATCATTGCCACA 561
DB 181 ThrPhePheLeuTyValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200

US-09-470-253-5
; Sequence 5, Application US/09470253
; Patent No. 6365371
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/09/470,253
; CURRENT FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-09-470-253-5

Alignment Scores:
Pred. No.: 9.06e-110 Length: 377
Score: 1054.50 Matches: 205
Percent Similarity: 77.44% Conservativeness: 49
Best Local Similarity: 62.50% Mismatches: 61
Query Match: 58.49% Indels: 13
DB: 3 Gaps: 2

US-10-089-688-1 (1-1053) x US-09-470-253-5 (1-377)

QY 1 ATGCCGTTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGGAAGATCTGAAG 60
DB 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20
QY 61 GAGACATGGCTGTCGAAAG-----CAAGAC 90
DB 21 GluValLeuThrIleLeuAspLysLeuProProLysLysLeuAspGlyAsnIle 40
QY 91 ATTTCTGATAAAAGCAGAAAGCTACAGAAAGTTTCAAAATCTCGTTGCCATG 150
DB 41 GlnSerAspLysLysTyAspLysAlaLeuAspGluValSerLysAsnValAlaMetIle 60
QY 151 AAGAAATTCGTATGGCACAATGAAAGAGCTTCAGACAGAA-----GCAGTA 201
DB 61 LysSerPheIleTyArgSerAlaGluProSerSerGluHisValValGlnVal 80
QY 202 GCTCAACTGCTCAGAACTCTATAATAGTGGCTCTTAGCACCCTGTTAGCTGATTTA 261
DB 81 AlaGlnLeuAlaGlnGluValTyArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 100
QY 262 CAGCTCATTGACTTTGAGCGCAAAAGACGTGGCTCAAAATTTCAACAATATTTCTCAGA 321
DB 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnLeuLeuArg 120
QY 322 AGACAAATGGTACGAGAACTCTACTGTTGAATACATCTGCACCACCAAGAAATTTTG 381
DB 121 ArgGlnIleGlyThrArgSerProThrValGluTyLeuGlyAlaArgProGluIleLeu 140
QY 382 TTCATGTTATTAAGAGGTGAATCTCCAGAAATAGCTCTAAATTTGGGAATATGTTA 441
DB 141 IleGlnLeuValGlnGlyTySerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160
QY 442 AGAAGATGATCAGACATGAACCACTTGCAGAAATCATTTGTTGTCGGAACAGTTTAT 501
DB 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTySerAspValPheTy 180
QY 502 GATTCTTCAGATATGTCGAATGTCAACATTTTGACATAGCTTCAGATCATTGCCACA 561
DB 181 ThrPhePheLeuTyValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200
```


Db	41	GlnSerAspIysIbIysTyrsbIysIalaleuaspGluValSerIysAsnValalaMetIle	60
Qy	151	AAAGAAATTCGTATGGCAACAATGAAAGAGCCCTCAGACAGAA-----GCAGTA	201
Db	61	LYSerPheIleTyGlyAsnAspSerAlaGluProSerSerGluHisValGlnVal	80
Qy	202	GCTCAACTTCCTCAGAACTCTATAATAGTGGGCTCTTAGCACCCCTGGTAGCTGATTTA	261
Db	81	AlaGlnLeuAlaGlnIuValTyrsAsnAlaAsnIleLeuProMetLeuIleIysMetLeu	100
Qy	262	CAGCTCATTTCTTGAGGGCAAAAAGACGCTGAGTCTCAAAATTTTCAACAATATTTCTCAGA	321
Db	101	ProLysPheGluPheGluCysIysIysAspValGlyGlnIlePheAsnAsnLeuLeuArg	120
Qy	322	AGACAAATTTGGTAGAGAACTCTACTGTTGAATACATCTGCACCCACAGAATATTTTG	381
Db	121	ArgGlnIleGlyThrArgSerProThrValGluTyrlleuGlyAlaArgProGluIleLeu	140
Qy	382	TTCATCTGTAATCAAGGGGTATGAATCTCCAGAAATAGCTCTAAATCTGTGGAATAAGTTTA	441
Db	141	IleGlnLeuValGlnGlyTyrsValProAspIleAlaLeuThrCysGlyIleuMetLeu	160
Qy	442	AGAGAAATGCATCAGACATGAACCACTTCGCAAAAATCATTTTGGTGGCGAACAGTTTAT	501
Db	161	ArgGluSerIleArgHisAspHisLeuAlaIysIleIleLeuTyrsAspValPheTyrl	180
Qy	502	GATTCTTCAGATATGTGCAATGTCAACATTTGCACATTTGCATGCTCAGATGCAATGGCCACA	561
Db	181	ThrPhePheLeuTyValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr	200
Qy	562	TTCAAGGATTTACTTACAGACATAAATGTGCTAGTGCAGATAATTTTGGAAACAGCATTA	621
Db	201	PheLysGluLeuThrThrArgHisIysAlaIleIleAlaGluPheLeuAspSerAsnTyrl	220
Qy	622	GATAGATTTTTCAGTGAATATGAGAAAGTTACTTCATTCAGAAATATATGTGACAAAAAGA	681
Db	221	AspThrPheAlaGlnTyrlGlnAsnLeuLeuAsnSerIysAsnTyrlValThrArgArg	240
Qy	682	CAGTCACTCAAGCTTCTCGGTGAACACTACTACTAGATAGACACACTTCACAATATAGCA	741
Db	241	GlnSerLeuIysIleuLeuGlyIuIleuLeuLeuAspArgHisAsnPheAsnThrMetThr	260
Qy	742	AAATACATCAGTAAACCTCGAGAACCTCAAAATTAATGATGAACCTGCTCGCAGACAAAAGT	801
Db	261	LysTyrlIleSerAsnProAspAsnLeuArgLeuMetMetGluLeuLeuArgAspIysSer	280
Qy	802	CGCAACATCCAGTTTGAGCCCTTTCAGCTTTTAAAGGTGTTGTAGCCAAATCTCTAACAG	861
Db	281	ArgAsnIleGlnTyrlGluAlaPheHisValPheIysValPheValAlaAsnProAsnLys	300
Qy	862	AGCGACCCATCTTAGACATCTCTCTCAAGAACCCAGCCCAACTCATAGAGTTTCTTCAGC	921
Db	301	ProLysProIleSerAspIleuAsnArgAsnArgGluIysIleuValGluPheLeuSer	320
Qy	922	AAGTTTCAAGACGACGAGCGGAT	945
Db	321	GluPheHisAsnAspArgThrAsp	328

RESULT. T 9

```

US-09-248-796A-14303
; Sequence 14303, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: FOR DIAGNOSTICS AND T
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIORITY APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIORITY APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13

```

```

; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14303
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-14303

Alignment Scores:
Pred. No.: 1,28e-31
Score: 358.50
Percent Similarity: 70.27%
Best Local Similarity: 48.65%
Query Match: 19.88%
DB: 4
Length: 165
Matches: 72
Conservative: 32
Mismatch: 43
Indels: 1
Gaps: 1

```

US-10-089-688-1 (1-1053) x US-09-248-796A-14303 (1-165)

286	QY	AAAGACGTGGCTCAAAATTTTCAACAATAATTTCTAGAAGACAAATTTGGTACGAGAACTCCT	345
13	Db	LyseAspValValIleLeuPheLeuThrLeuLeuArgThrMetAlaAsnLysSerPro	32
346	QY	ACTGTTGAAATACATCTGC---ACCAACACAGAAATTTTGTTCATGCTTAATTTGAAAGGGTAT	402
33	Db	ThrValAspTyrLeuValHisSerLysProGluIleIleThrMetLeuIleLysGlyPro	52
403	QY	GAATCTCCAGAAATAGCTCTAAATTTGTGGAAATAATGTTAAAGAGAAATGCATGACATGAA	462
53	Db	GluAsnLeuGluIleGlyLeuIleCysGlyGlnIleLeuArgAspCysIleLysPheGlu	72
463	QY	CCACTTCGAAAATCATTTTGTGTGCGAACAGTTTTATGATTTCTTCAGATATATGTCGA	522
73	Db	ValIleAsnLysPheValLeuTyrSerProSerPheTyrAsnPhePheLysTyrValGln	92
523	QY	ATGTCAACATTTGACATAGCTTCAGATGTCATTTGCCACATTTCAAGGATTTTACTTACAAGA	582
93	Db	IleProThrPheAspIleAlaThrAspAlaMetThrLeuHisGluLeuLeuThrThr	112
583	QY	CATAAATGCTCAGTGCAGAAATTTTGTGGACAGCATTTATGATAGATTTTTCAGTGAATAT	642
113	Db	HisArgLysLeuValSerGluPheLeuGlyAsnAsnTyrAspValPheIleThrAlaIle	132
643	QY	GAGAAGTTACTTCATTCCAGAAAAATTATGTGCACAAAAAGACAGTCACCTCAAGCTTCTCGGT	702
133	Db	AsnLysLeuValThrSerLysAsnTyrValThrLysArgLeuSerValLysLeuLeuAsp	152
703	QY	GAACTACTACTAGATGACACAAC	726
153	Db	GluLeuValSerGlnArgSerAsn	160

RESULT 10

```

RESUL 10
US-09-248-796A-15202
; Sequence 15202, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AM
; TITLE OF INVENTION: FOR DIAGNOSTICS AN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,7
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15202
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15202

Alignment Scores:
Pred. No.: 5,15e-10
Len

```

Score: 166.00 Matches: 38
Percent Similarity: 59.46% Conservative: 28
Best Local Similarity: 34.23% Mismatches: 43
Query Match: 9.21% Indels: 2
DB: 4 Gaps: 1

US-10-089-688-1 (1-1053) x US-09-248-796A-15202 (1-125)

QY 1 ATCCGCTTCGGTGTGGAGAGTCTCAAAATCCTCAGACACATTTGTGAAGAATCTGAAG 60
||| ||| ||| :||| :||| :||| :||| :|||
Db 16 MetAlaPheLeuPheIysArgAsnProLysThrProGluLeuValArgAlaLeuAsn 35
||| ||| ||| :||| :||| :||| :||| :|||
QY 61 GAGAGCATGCTCTCTTGAAAAGCAAGACATTCTTGATAAAAAAGCACAAAAGGCTACA 120
::: ||| :||| :||| :||| :||| :||| :||| :|||
Db 36 AspGln-----ValLeuIysLeuAspTyrAlaSerProAspAsnAlaIysIysTyrGln 53
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 121 GAAGAAGTTTCCAAAATCTGGTGGCATGAAGAAGAAATCTGTATGGCACAAATGA AAAA 180
||| ||| ||| :||| :||| :||| :||| :||| :|||
Db 54 AspGluCysAlaArgTyrLeuLysAsnMetLysValIleLeuHisGlyAspAspGluVal 73
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 181 GAGCTCAGACAGACAGTAGCTCAACTGTGCTCAAGAACTCTATAATAGTGGGCTCCTT 240
||| ||| ||| :||| :||| :||| :||| :||| :|||
Db 74 GluProGlnProAspGlnIleThrGlnLeuAlaGlnIleIysSerThrAspCysLeu 93
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 241 AGCACCTGGTAGCTGATTTACAGCTCATTTGAGGGCGAAAAGAACGCTGGCTCAA 300
||| ||| ||| :||| :||| :||| :||| :||| :|||
Db 94 TyrTyrLeuValValAsnLeuArgLysLeuAspPheAspSerArgGluArgCysCysTyr 113
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 301 ATTTCACAAATATCTCAGAGACAAATTTGGT 333
||| ||| ||| :||| :||| :||| :||| :||| :|||
Db 114 IleValSerAspIleIleAlaThrAsnTyrGly 124
||| ||| ||| :||| :||| :||| :||| :||| :|||

RESULT 11

US-09-248-796A-15201
; Sequence 15201, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107195.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15201
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15201

Alignment Scores:

Pred. No.: 90 Length: 90
Score: 163.00 Matches: 31
Percent Similarity: 63.93% Conservative: 8
Best Local Similarity: 50.82% Mismatches: 22
Query Match: 9.04% Indels: 0
DB: 4 Gaps: 0

US-10-089-688-1 (1-1053) x US-09-248-796A-15201 (1-90)

QY 745 TATCATCAGTAAPAACTCGAGAACCTCAAAATTAATGATGAACCTGCTCGAGACAAAAGTCGC 804
::: ||| ||| ||| :||| :||| :||| :||| :||| :|||
Db 1 PhePheAspAlaAsnAsnLeuLysLeuThrMetLeuLeuSerAspLysLeuLys 20
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 805 ACATCCAGTTTGAGGCCCTTTCACGTTTTTAAAGGTGTTTGTAGCCAATCTCTAAACAAGACG 864
||||| ||| :||| :||| :||| :||| :||| :||| :|||
Db 21 AsnLeuGlnLeuGluGlyPheHisThrLeuLysPhePheValAlaAsnProLysArgSer 40
||| ||| ||| :||| :||| :||| :||| :||| :|||
QY 865 CAGCCCATCTAGACATCTCTCTCAAGAACCCAGGCAAACTCATAGTTCCTCAGCAAG 924
||| ||| ||| :||| :||| :||| :||| :||| :|||

Db 264 AspGluLeuLysGluLysThrLysGlnPheAspAap-----SerLysLysLys 279
Qy 529 ACATTGACATAGCTTCAGATGCATTTGCCACATTCAGGATTTACTTACAGACATATAA 588
Db 280 LeuThrGluLeuGluAsnAspLeuThrSerThrLysLysGluLeuGluThr---GluLys 298
Qy 589 TTGCTCAGTCAGCAATTT-----TTGCAACAGCATTTATGATAGATTTTTCAGTGAATAT 642
Db 299 ThrGlnThrSerLysPheLysAsnLeuGluGluArgLysAspLysGluLeuValLysLeu 318
Qy 643 GAGAAG-----TTACTTCATTCAGAAATATTATGTACAAAAGACAGCTCACTCAAG 693
Db 319 AsnLysGluLeuGluLeuLysAsnAspAsnSerGlyAlaLysLysGluLeuGlu 338
Qy 694 CTTCTCGGTGAACTACTACTAGATAGACACAACTTCACAAATATGACAAAATACATCAGT 753
Db 339 LysValSerLysLeu-----GluSerGluLeuGluLeuSerLysGluLeuGlu 355
Qy 754 AAACCTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAGTCGCAACATCCAG 813
Db 356 AspLysLysSerValMetLysGlnHisAspGluLeuLysGluGlnThrLysGluLysAsn 375
Qy 814 TTGAGGCTTTTCACTGCTTCAAGAACCGCCMAACTCATAGATTCCTCAGCAAGTTTCAG--- 930
Db 392 LeuAspGluLeuGlnLysGluLeuAspAlaAlaLeuSerPheLysAspLysPheGluThr 411
Qy 931 --AACGACAGCGGATTTGATGACGAGCTTCCGTACCGACGACGAATTCGCGGTGAT 987
Db 412 AlaSerAlaLysLeuThrGlnSerThrSerAspLeuGluAlaAlaAsnLysLysLeuAsn 431
Qy 988 TTACGCTTAAACCGCGAGCGCTGGGATCAGGATTTGAAGAGACACCGCTCAGCAAGAA 1047
Db 432 IleLeuLysSerGluLysGluLysThrGluGlnGluLeuGluLysLeuThrLysGlnHis 451
Qy 1048 GCT 1050
Db 452 Ala 452
RESULT 13
US-09-134-000C-5420
; Sequence 5420, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5420
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5420
Alignment Scores:
Pred. No.: 0.000946 Length: 541
Score: 113.00 Matches: 89
Percent Similarity: 38.76% Conservative: 61
Best Local Similarity: 23.00% Mismatches: 141
Query Match: 6.27% Indels: 96
DB: 4 Gaps: 21
US-10-089-688-1 (1-1053) x US-09-134-000C-5420 (1-541)

Qy 40 GACATTCTGAGAAATCTGAAGAGAGCATGCTGCTTCTGGAAGAAAGCAA---GACATTTCT 96
Db 173 AspIleIleLys---LeuProPheAlaLeuLysMetLeuAspLysHisLeuGluIleSer 191
Qy 97 GAT-----AAAAAAGCAGAAAGAGGCTACAGAAAGAA 126
Db 192 AspGlyAspLeuProHisAspAlaIleValLeuSerArgGluLysLeuLysAsnSerGlu 211
Qy 127 GTTTCAAAAAATCTGGTTGCCATGAAAGAAATCTGTATGGCACAACAAATGAAAAAGAGCCT 186
Db 212 PheThrLysAsnAspTyrSerLeuLeuThrLeuTyrSerHisAlaAsnProLysLys--- 230
Qy 187 CAGACAGACAGCAGTAGCTCAACTGCTCAAGAA-----CTCTATATAGTAGGGCTCCTT 240
Db 231 -----GlnLeuAlaLeuGluValSerIlePheAsnSerTyrSerThr 244
Qy 241 AGCACCCCTG-----GTAGCTGATTTTACAGCTCATTTGAC 273
Db 245 AsnCysLeuLeuLeuValTyrLeuGlnAspGluLysValLysGluThrAspLeuValLe 264
Qy 274 TTGAGGGCAAAAAGACGCTGGCTCAA---ATTTCACAC-----AATATTCTCAGAAAGA 324
Db 265 IleGluAsnAlaIleAspValLeuGlnGluLysPheAsnThrGluAsnLeuLysLys 284
Qy 325 CAAATTGGTACGAGAACTCCTCTGTTGAATACATCTGCACCCCAAGAAATATTTTGTTC 384
Db 285 GluArgTyrThrArgLeu-----AsnAsnLeuAlaAsp 295
Qy 385 ATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGCAATATGTTAAGA 444
Db 296 AlaIleLeuGln-----AsnThrProGlnAsnProAspGluLeuAsnSerLeuLeuArg 313
Qy 445 GAATGATCATCAGCATGAACCACTTGCAAAAATCATTTTGTGTCGGAACACATTTTATGAT 504
Db 314 GluValGlnMetGlnGlu-----LeuAsp 321
Qy 505 TTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGATTTGCCACATTC 564
Db 322 SerTyrGlnAlaIleAlaPheSerThrLysAspThrAsnThrGlnLeuMetLysGluArg 341
Qy 565 AAGGATTTTACTTACAAGACATAAATGCTCAGTGCAGAAATTTTGGAAACAG---CATTAT 621
Db 342 IleIleAlaLeuLeuArgThrLeuArgValLysSerIlePheAspGlnLeuAsnTyr 361
Qy 622 GATAGATTTTTCAGTGAATATGAGAAGTACTTTCATTTCAGAAAATATTTGTCACAAAAGA 681
Db 362 SerAlaValLeuPheAsnPheAsnGlu-----SerAspGlyLysIleThrLysLeu 378
Qy 682 CAGTCACTGAAGCTTCTCGGTGAATCTACTACTAGATAGACACAACTTCACAAATATGACA 741
Db 379 GlnLeuSerArgLeuLeuAlaGluLeuGluAsnAspThrLeuThrValAlaVal 398
Qy 742 AATATCATCAGTAAACCTGAGAACCTCAAAATTAATG-----ATGAACCTGCTG 789
Db 399 SerSerLeuLysSerArgGluGlyIleLysGluLeuLeuLysCysLeuAspIleLeu 418
Qy 790 CGAGACAAAAGTCGCAACATCCAGTTTGGAGCCCTTTCAC-----GAGGCTGCTG 828
Db 419 ArgPheAsn-----GluThrPheTyrAsnGlyProIleValThrLeu 432
Qy 829 -----GTTTTTAAGGTGTTTGTGTA-----GCCAATCTCAACAAG 861
Db 433 AlaAspIleGlyValPheLysAsnPheIleArgGluAspGlnLeuGluAsnLeuAspGlu 452
Qy 862 ACGCAGCCC---ATCCTAGACATCCTCCTCAAGAACCGAGCCCAAACTCATAGATTTCTC 918
Db 453 LeuIleProArgAlaLeuTyrGlnLeuAlaGluAsnAsnTyrAspLeuPheGluThrLeu 472
Qy 919 ---AGCAAGTTTTCAGAACGACAGCGGATTTGATGAGCAGGTTTCCGTACCGACGACGAAT 975
Db 473 TyrSerPhePheGlnAsnAsnArgAsnTyrLysGlnThrSerGluAlaMetPheLeuHis 492
Qy 976 TCCCGGGTCGATTTACGCGTTAAACCGCGAGCGGTGGGATCAGG----- 1020

```
Db 493 SerLysThr-----IleArgTyrArgLeuAsnLysValGluGlnLeuLeuAspIle 509
QY 1021 GATTTGAAGAGACACGCTCAG 1041
Db 510 AspLeuAlaAsnProLeuGln 516

RESULT 14
US-09-417-485D-6
; Sequence 6, Application US/09417485D
; Patent No. 6541202
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/09/417,485D
; CURRENT FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)..(335)
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-09-417-485D-6

Alignment Scores:
Pred. No.: 0.00183 Length: 2184
Score: 113.00 Matches: 68
Percent Similarity: 40.94% Conservative: 54
Best Local Similarity: 22.82% Mismatches: 98
Query Match: 6.27% Indels: 78
DB: 4 Gaps: 18

US-10-089-688-1 (1-1053) x US-09-417-485D-6 (1-2184)
QY 265 CTCATTGACTTGGGCAAAAAGACGTGCTCAAAATTTTCAACAAT-----312
Db 401 LeuLeuAsnTyrIlePheLysSerAspGlnThrTyrPheTyrHisAsnAsnPheIleAsp 420
QY 313 -----ATTCTCAGAAGACAAAT-----GGTAGCAGAACTCTCTACTGTTGAA 354
Db 421 GluTyrLysGlnLysIleCysLysGlnIleLysCysSerThrLysLysAsnAspIleSer 440
QY 355 TACATCTGCACC-----CAACAGAAATTTTGTTCATGTTATTGAAA-----GGG 399
Db 441 HisIleIleThrSerArgLysGluAsnHisLeuPheHisValGlnLysLeuGluAsnAsn 460
QY 400 TATGAATCTCCAGAATAGCTTAATTTGCGAATAAGTTAAGAGATGATCAGACAT 459
Db 461 TyrLysHisProAsnIle-----AsnLysGlnLeuArgLys 472
QY 460 GAACCACTTGCAAAAATCATTTGTGTGGCAACAGTTTATGATTTCTTCAGATATGTC 519
Db 473 ThrLysIleLeuLysTyrValTyr-----AsnTyrPheLysGluPheIleAsnVal 490
QY 520 GAAATGTCAACATTTGACATAGCTTCAGATGCAATTTGCCACATTCACAGGATTTACTTACA 579
Db 491 IleAsnThrLysPheGlyLysIleTyrArgLysPhePheProArgLysHisIleLeuAsn 510
QY 580 AGA---CATAATG-----591
Db 511 LysIleHisIlePheLysIleIleArgLeuGlnIleIleLysLysTyrArgIleIle 530
QY 592 -----CTCAGTGCAGAAATTTTGGAAACAGCAT---TATGATAGATTTTTCAGTGAA 639
Db 531 AsnIleArgMetAsnArgLysPheIleLysGlnLysValTyrAspThrPhePheLysAsn 550
```

```
QY 640 TATGAG-----AAGTTACTTTCATTGAGAAATATTATG 672
Db 551 TyrAspPheLeuSerPheSerPheLysThrTyrLysIleIleAsnPheMetValTyrIle 570
QY 673 ACAAAGA---CAGTACTGAAGCTTCGGTGAACACTACTAGATAGACACAACTTC 729
Db 571 ThrLysLysCysIleProIleLysLeuLeuGly-----SerLysHisAsnPhe 586
QY 730 ACATTTATGACAAAATACATCATGATAA-----CCTCAGAACCTCAAA 771
Db 587 LysIlePheLeuLysAsnValLysLysPheLeuLeuPheAsnTyrLysGluSerPheSer 606
QY 772 TTAATGATGAACCTGCTGCAGACAAAGTCGCAACATCCAGTTTCAGGCTTTTCACGT 831
Db 607 LeuAsnGlnValMetLysAsnIleLysValLysAsnIle---PheGlnLys---LysIle 624
QY 832 TTTAAGGTGTTTGTAGCCAATCCTAAAGAGCGACGCCATCTCTAGAC-----879
Db 625 SerLysTyrAsnIleLysAsnArgIleLeuLeuLysAsnIlePheAspAsnAsnTyrGlu 644
QY 880 -----ATCTCTCTCAAGACACGAGCCAAACTCATAGAGTTCTCTCAGC-----AAG 924
Db 645 AsnLysIleLeuHisArgAsnAsnLysGluIleIleThrAsnIleAsnAspAsnIleLys 664
QY 925 TTTCAGAACGACAGACGAGGATTTGTATGATGAGCAGTTTC-----GTACCGACGACG 972
Db 665 IleTyrAsnLysLysAsnAspAsnLeuAsnAsnSerPheLysIleLysThrThr 682

RESULT 15
US-09-134-001C-4053
; Sequence 4053, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4053
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4053

Alignment Scores:
Pred. No.: 0.0018 Length: 533
Score: 110.50 Matches: 73
Percent Similarity: 40.00% Conservative: 57
Best Local Similarity: 22.46% Mismatches: 108
Query Match: 6.13% Indels: 87
DB: 3 Gaps: 18

US-10-089-688-1 (1-1053) x US-09-134-001C-4053 (1-533)
QY 40 GACATTGTGAAGAATCTGAGAGAGCATGCTGTTCTG-----GAAAGCA 87
Db 237 AspIleLeuArgGlnPheSerAspSerIleAspValLeuIleThrGlyHisGlnHisArg 256
QY 88 GACATTTCTGATAAAGCAAGCAAGCT---ACAGAGAGAGTTTCCAAAATCTGTT 144
Db 257 AspIleAlaThrIleLysAsnGlnThrAlaIleIleGlnProGlySerLysGlyThrLys 276
QY 145 GCCATGAAGAAATTTCTGATGGCACAATCAAAA-----180
Db 277 ValGlyLysIleValIleGluTyrThrHisAspLysLysValLeuLysGluCysAsn 296
```

```
Qy 181 -----GAGCTCTCAGACAGAGCAGTAGCTCAACTT 210
Db 297 LeuMetAsnValHisAsnSerThrSerPheLysProAsnAspGluAspIleAala---Leu 315
Qy 211 GCTCAAGAACCTCTATATAGTGGCTCCTTAGCACCTGGTAGCTGATTTACAGCTCATT 270
Db 316 ArgAsnGlnLeu-----GluAspTrpLeuAspThrGlnIleAlaGluLeuProTyrAla 333
Qy 271 -----GACTTTGAGGGCAAAAAGAC-----GTGGCTCAAAATTTTC 306
Db 334 MetArgIleAsnAsnSerPheGluAlaArgLysSerProHisAlaPheValAsnLeuLeu 353
Qy 307 AACATATATTCAGAGACAAATTTGGTAGAGAACTCTACTGTTGAATACATCTGCACC 366
Db 354 AsnTyrIleLeuLeuGluLysSerGly-----AlaAspIleAalaCysThr 368
Qy 367 CAACAGAAATATTTGTTTCATGCTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAAT 426
Db 369 -----AlaLeuPheAspSerAlaAsnGlyPheAspGluLys----- 380
Qy 427 TGTGGAATAATGTTAAGAGAAATGCATCAGACATGAAACCACTTGCAGAAATCAATTTGTGG 486
Db 381 -----ValThrMetArgAspIleIleAsnAsnTyrPro----- 391
Qy 487 TCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATA----- 540
Db 392 -----PheProAsnThrPheLysValIleGluLeuSerGlyLysAspIleLysLeu 408
Qy 541 GCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAGACATAAATTCCTCAGTGCA 600
Db 409 AlaIleGluArgSerAlaSerTyrPheAspIleValAsnHisLysIleThrValAsnLys 428
Qy 601 GAATTTTGGAA-----CAGCATTTATGAT---AGATTTTTCAGTGAATATGAGAAG 648
Db 429 GluPheLeuGluProLysProGlnHisPheAsnTyrAspIlePheAlaGlyIleGlnTyr 448
Qy 649 TTACTTCATTCAGAAAATATGTGACAAAAGACAGCTCAGTGAAGCTTCTCGGTGAACATA 708
Db 449 ThrIleHisValSerHisProTyrGlyGluArgValSerAspLeuLeuIleAsnAspAla 468
Qy 709 CTACTAGATAGACACAAC---TTCACAATTTATGACAAA----- 744
Db 469 ProLeuGlnSerAspGlnIleTyrThrIleCysValAsnAsnTyrArgAlaValGlyGly 488
Qy 745 -----TACATCAGTAAACCT-----GAGAACCTC 768
Db 489 GlyAsnTyrAspMetTyrValAsnLysProValIleLysAspIleGlnIleGluGlyAla 508
Qy 769 AATTAATGATGAACCTGCTG-----CGAGACAAAAGTCGCAACATCCAGTTTGAGGCC 822
Db 509 GlnLeuLeuIleAspTyrLeuSerHisAsnAspLeuSerGlnIleProGlnValIleAsp 528
Qy 823 TTTCACGCTTTTAAAG 837
Db 529 PheAsnValValLys 533
```

Search completed: November 10, 2004, 20:54:02
Job time : 47.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:43:16 ; Search time 165.5 Seconds
(without alignments)
4494.826 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atccgttcctccgtttgggaa.....cagctcagcagaagcttaa 1053

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 3133240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10089688/runat_10112004_152249_4931/app_query.fasta_1.1223
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOPCCI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=15 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10089688 @CGN 1 1 199 @runat_10112004_152249_4931
-NCPU=6 -ICPU=3 -NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

* Query
Result

No.	Score	Match	Length	DB	ID	Description
1	1635.5	90.7	341	16	US-10-408-765A-911	Sequence 911, App
2	1635.5	90.7	496	14	US-10-239-079-5	Sequence 5, Appli
3	1635.5	90.7	552	14	US-10-239-079-6	Sequence 6, Appli
4	1614.5	89.5	341	14	US-10-025-730-3	Sequence 3, Appli
5	1297.5	72.0	337	14	US-10-025-730-1	Sequence 1, Appli
6	1117	62.0	339	14	US-10-025-730-4	Sequence 4, Appli
7	1054.5	58.5	377	14	US-10-025-730-5	Sequence 5, Appli
8	750	41.6	446	15	US-10-425-114-52177	Sequence 52177, A
9	747.5	41.5	337	17	US-10-425-114-322057	Sequence 162057, A
10	747	41.4	336	16	US-10-437-963-161075	Sequence 161075, A
11	741	41.1	412	15	US-10-425-114-54669	Sequence 54669, A
12	721.5	40.0	339	15	US-10-424-599-224185	Sequence 224185, A
13	652	36.2	336	17	US-10-425-115-201848	Sequence 201848, A
14	649	36.0	351	15	US-10-425-114-46025	Sequence 46025, A
15	639.5	35.5	337	16	US-10-437-963-183779	Sequence 183779, A
16	635.5	35.2	342	15	US-10-425-114-48872	Sequence 48872, A
17	635.5	35.2	363	17	US-10-425-115-215481	Sequence 215481, A
18	593	32.9	334	16	US-10-437-963-146612	Sequence 146612, A
19	501	27.8	337	15	US-10-424-599-209662	Sequence 209662, A
20	459.5	25.5	256	17	US-10-425-115-322099	Sequence 322099, A
21	400.5	22.2	197	16	US-10-767-701-37391	Sequence 37391, A
22	387	21.5	274	17	US-10-425-115-233795	Sequence 233795, A
23	379.5	21.0	173	15	US-10-424-599-158029	Sequence 158029, A
24	362	20.1	205	17	US-10-425-115-201847	Sequence 201847, A
25	353.5	19.6	140	16	US-10-767-701-42587	Sequence 42587, A
26	322	17.9	351	17	US-10-425-115-201850	Sequence 201850, A
27	321	17.8	184	15	US-10-424-599-149595	Sequence 149595, A
28	286.5	15.9	294	16	US-10-437-963-119830	Sequence 119830, A
29	255	14.1	149	17	US-10-425-113-350620	Sequence 350620, A
30	255	14.1	161	15	US-10-425-114-70029	Sequence 70029, A
31	218	12.1	90	17	US-10-425-115-358599	Sequence 358599, A
32	209	11.6	135	11	US-09-864-408A-6108	Sequence 6108, Ap
33	178	9.9	94	17	US-10-425-115-282537	Sequence 282537, A
34	168	9.3	83	15	US-10-425-114-44450	Sequence 44450, A
35	135.5	7.5	116	15	US-10-424-599-210348	Sequence 210348, A
36	126	7.0	1182	15	US-10-282-122A-53445	Sequence 53445, A
37	125.5	7.0	74	15	US-10-425-114-56749	Sequence 56749, A
38	116.5	6.5	888	10	US-09-893-519A-73	Sequence 73, Appli
39	113	6.3	2184	14	US-10-304-095-6	Sequence 6, Appli
40	111.5	6.2	70	15	US-10-424-599-210351	Sequence 210351, A
41	111.5	6.2	790	16	US-10-408-765A-2178	Sequence 2178, Ap
42	110.5	6.1	797	16	US-10-437-963-163096	Sequence 163096, A
43	110.5	6.1	4131	14	US-10-369-493-5136	Sequence 5136, Ap
44	109.5	6.1	367	15	US-10-282-122A-45776	Sequence 45776, A
45	109.5	6.1	564	15	US-10-282-122A-61170	Sequence 61170, A

ALIGNMENTS

RESULT 1
US-10-408-765A-911
; Sequence 911, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660888.465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 341
; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-408-765A-911

Alignment Scores:
Pred. No.: 3,18e-151 Length: 341
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 16 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-408-765A-911 (1-341)
QY 1 ATGCGCTCCCGTTGGAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 1 MetProPheProPheGlySerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAAGCAGAAAGGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAGTTTCCAAAATCTGGTGGCCATGGAAGAAATCTGTATGCGCACAAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCTCAGCAGACAGTAGCTCAACTGCTCAAGAACTCTATAATAGTGGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTAGCTGATTATTCAGCTCATTCAGCTTTGAGGGCAAAAGAGCGTCA 300
DB 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACATATTTCTCAGAAAGCAATTTGGTACGAGAACTCTACTGTGTGAATCATC 360
DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCCAACAGATATTTGTTGATGTTTAAAGGATGATGATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAATTTGGAATAATTTAAGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGCGTCGGAACAGCTTTTATGATTTCTCAGATATGTCGAAATCTCAACATTTGACATA 540
DB 161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCAATTTGCCACATTTCAAGGATTTACTTACAAGACATAAAATTTGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAATTTTGGACACAGCATATGATAGATTTTTCAGTGAATATGAGAGATTTACTTCATCA 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATTTATGTCACAAAAGACAGTCACTGAGCTTCTCGGTGAACTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAATTTCAACATTTATGACAAATACATCAGTAAACCTTGAAACCTCAAAATTTATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACCTGCTGCGAGACAAAGTCGCAACATTCAGTTTGGAGCCCTTTCACGTTTTTAAGGTG 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCAAATCTTAAACAGACGCGCCATCTCTAGACATCTCTCAAGAACACAGGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300

; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: ANIC-BP1-ligand
; FILE REFERENCE: ANIC-BP-1-ligand
; CURRENT APPLICATION NUMBER: US/10/239,079
; CURRENT FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gal4-ANIC-BP-1
; OTHER INFORMATION: fusion protein
US-10-239-079-5

Alignment Scores:
Pred. No.: 3,52e-151 Length: 496
Score: 1635.50 Matches: 329
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 11
Query Match: 90.71% Indels: 9
DB: 14 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-239-079-5 (1-496)
QY 1 ATGCGCTCCCGTTGGAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 156 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 175
QY 61 GAGAGCATGGCTGTTCTGGAAGCAAGACATTTCTGATAAAAGCAGAAAGGCTACA 120
DB 176 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 195
QY 121 GAAGAGTTTCCAAAATCTGGTGGCCATGGAAGAAATTTGTATGCGCACAAATGAAAA 180
DB 196 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 215
QY 181 GAGCTCAGCAGACAGTAGCTCAACTGCTCAGAACTCTATATAGTGGCTCCTT 240
DB 216 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLeuTyrAsnSerGlyLeuLeu 235
QY 241 AGCACCTGGTAGCTGATTATTCAGCTCATTCAGCTTTGAGGGCAAAAGAGCGTGGCTCAA 300
DB 236 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 255
QY 301 ATTTTCAACATATTTCTCAGAAAGCAATTTGGTACGAGAACTCTACTGTGTGAATCATC 360
DB 256 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 275
QY 361 TGCACCCCAACAGATATTTGTTGATGTTTAAAGGATGATGATCTCCAGAAATAGCT 420
DB 276 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 295
QY 421 CTAATTTGGAATAATTTAAGAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 480
```

Db 296 LeuAanCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 315
Qy 481 TTGTGCTCGGAACAGTTTATGATTTCTTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 316 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 335
Qy 541 GCTTCAGATGCTATGTCACATTCAGAGTATCTTCAAGACATATAATGCTCAGTGCA 600
Db 336 AlaserAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 355
Qy 601 GAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAGTTTACTTCATTCA 660
Db 356 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 375
Qy 661 GAAATATATGTCACAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGA 720
Db 376 GluAenTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 395
Qy 721 CACAACCTTCACAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATCATG 780
Db 396 HisAenPheThrIleMetThrLysTyrIleSerLysProGluAenLeuLysLeuMetMet 415
Qy 781 AACCTGCTGCAGACAAAAGTCGCAACATCCAGTTTTCAGGCTTTTCACGTTTTCAGGTG 840
Db 416 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 435
Qy 841 TTTGTAGCAATCTTAACAGACGAGCCCATCTAGACATCTCTCAAGAACACGAGCC 900
Db 436 PheValAlaAenProAenLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 455
Qy 901 AAATCTATAGATTCCTCAGCAAGTTTCAGAACACAGGACGAGCGGATTTGTATGAGCAGTCC 960
Db 456 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 470
Qy 961 GTACCGACGCAATTCCTCGGCTGATTTACGCTGTTAAACCGCGGCGCTGGGATCAGG 1020
Db 471 AspGluGlnPheAenAspGluLysThrTyrLeuValLysGln-----IleArg 486
Qy 1021 GATTGGAAGACAGCTCAGCAAGAGCT 1050
Db 487 AspLeuLysArgProAlaGlnGlnGluAla 496

RESULT 3

US-10-239-079-6

; Sequence 6, Application US/10239079

; Publication No. US2003014846A1

; GENERAL INFORMATION:

; APPLICANT: Merck Patent GmbH

; TITLE OF INVENTION: ANIC-BP1-ligand

; FILE REFERENCE: ANIC-BP-1-ligand

; CURRENT APPLICATION NUMBER: US/10/239,079

; CURRENT FILING DATE: 2002-09-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 552

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: LexA-ANIC-BP-1

; OTHER INFORMATION: fusion protein

US-10-239-079-6

Alignment Scores:

Pred. No.:	3,62e-151	Length:	552
Score:	1635.50	Matches:	329
Percent Similarity:	94.29%	Conservative:	1
Best Local Similarity:	94.00%	Mismatches:	11
Query Match:	90.71%	Indels:	9
DB:	14	Gaps:	2

US-10-089-688-1 (1-1053) x US-10-239-079-6 (1-552)

RESULT 4

Qy 1 ATGCGCTTCCGCTTTCGGAAGTCTCACAATCTCCAGCAGACATTGTGAAGATCTCAAG 60
Db 212 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAenLeuLys 231
Qy 61 GAGAGCATGCTGCTTCTCGGAAAAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACA 120
Db 232 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 251
Qy 121 GAGAGAGTTTCCAAAAATCTGGTGGCATGCAAGAAATCTGTATGGCACAATAATGAAAA 180
Db 252 GluGluValSerLysAenLeuValAlaMetLysGluIleLeuTyrGlyThrAenGluLys 271
Qy 181 GAGCCTCAGACAGACAGCAGTCTCACTTGTCTCAAGAACTCTATATAGTGGCTCTT 240
Db 272 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAenSerGlyLeuLeu 291
Qy 241 AGCACCTCTGTAGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGCTGGCTCAA 300
Db 292 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 311
Qy 301 ATTTTCAACAATATCTCAGAAAGCAAAATTTGGTACGAGAACTCTCTACTGTGTGAATCATC 360
Db 312 IlePheAenAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 331
Qy 361 TGCACCCACAGATATTTTGTTCATGTTATTCAGAGGTATCAATCTCCAGAAATAGCT 420
Db 332 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 351
Qy 421 CTAAATTTGCAATAATGTTAAAGAGATGTCATCAGACATCAACACACTTGCAAAATCATT 480
Db 352 LeuAenCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 371
Qy 481 TTGTGTCGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 372 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 391
Qy 541 GCTTCAGATGCTATTCGCCACATTCAGAGATTTACTTACAAGACATAAATGTCTCAGTGCA 600
Db 392 AlaserAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 411
Qy 601 GAATTTTGGAAACAGCATTTATGATATGATTTTTCAGTGAATATCAGAAAGTTTACTTCA 660
Db 412 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 431
Qy 661 GAAATATATGTGACAAAAAGACAGTCACTCAAGCTTCTCGGTGAACCTACTACTAGATAGA 720
Db 432 GluAenTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 451
Qy 721 CACAACCTTCACAATATGACAAAATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
Db 452 HisAenPheThrIleMetThrLysTyrIleSerLysProGluAenLeuLysLeuMetMet 471
Qy 781 AACCTGCTCGAGACAAAAGTGCACATCCAGTTTTCAGGCTTTTCACGTTTTCAGGCTG 840
Db 472 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 491
Qy 841 TTTGTAGCCCAATCTTAACAAAGACGAGCCCATCTTAGACATCTCTCTCAGAACCCAGGCC 900
Db 492 PheValAlaAenProAenLysThrGlnProIleLeuAspIleLeuLeuLysAenGlnAla 511
Qy 901 AAATCATAGATTTCTCAGCAAGTTTTCAGAACGACGACGAGCGATTTGTATGACAGTCC 960
Db 512 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu-----IleArg 526
Qy 961 GTACCGACGCAATTCCTCGGCTCGATTTACGCTGTTAAACCGCGGCGCTGGGATCAGG 1020
Db 527 AspGluGlnPheAenAspGluLysThrTyrLeuValLysGln-----IleArg 542
Qy 1021 GATTGGAAGACAGCTCAGCAAGAGCT 1050
Db 543 AspLeuLysArgProAlaGlnGlnGluAla 552

```
US-10-025-730-3
; Sequence 3, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE: -
; OTHER INFORMATION: 9262934
US-10-025-730-3

Alignment Scores:
Pred. No.: 3.64e-149 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 14 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-025-730-3 (1-341)

QY 1 ATGCGCTCCCGTTGGGAAGTCTCACAATCTCCAGCAGACATTTGTAAGAATCTGAAG 60
Db 1 MetProPheProPheGlyLysSerHisLysSerProAlaPheValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGGAAGACAGACATTTCTGATAAAGACAGAAAGGCTACA 120
Db 21 GluSerMetAlaValLeuGluLysGlnAspLysSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAGCTTCCAAAATCTGTTGCCATGAAGAAATCTGTATGGCACAATGAAAAA 180
Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluLysLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGCTCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTAGCTGATTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAA 300
Db 81 GlyThrLeuValAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValAlaGln 100
QY 301 ATTTTCAACAATATCTCAGAGACAAATTTGTAGAGAACTCTCTACTGTTTGAATACATC 360
Db 101 IlePheAsnAsnIleLeuLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCCAACAATATTTTGTTCATGTTATTGAAAGGCTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAATTTGGTAATATTTAAGAGAAATGTCATCAGACATGACCACTTGCAGAAATCATTT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGCGTCGGAACAGTTTTATCATTTCTTCAGATATGTCGAATGTCAACATTTGCACATA 540
Db 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGCAATTTGCCACATTCAGAGATTTACTTACAGACATATAAATTTGCTCAGTGA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200

US-10-025-730-1
; Sequence 1, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-10-025-730-1

Alignment Scores:
Pred. No.: 4.22e-118 Length: 337
Score: 1297.50 Matches: 263
Percent Similarity: 85.26% Conservative: 32
Best Local Similarity: 76.01% Mismatches: 38
Query Match: 71.96% Indels: 13
DB: 14 Gaps: 3

US-10-089-688-1 (1-1053) x US-10-025-730-1 (1-337)

QY 1 ATGCGCTCCCGTTGGGAAGTCTCACAATCTCCAGCAGACATTTGTAAGAATCTGAAG 60
Db 4 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys 22
QY 61 GAGAGCATGGCTGTTCTGGAAGACAGACATTTCTGTATAAAGACAGAAAGGCTACA 120
```

```
Db 23 AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer 39
QY 121 GAAGAGTTTCCAAAATCTGGTCCCATGAAAGAAATCTGTATGGCACAAATGAAAAA 180
Db 40 GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys 59
QY 181 GAGCCTCAGACAGAGTAGTACCTCAACTGCTCAAGAACTCTATAAGTGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu 79
QY 241 AGCACCTGGTAGCTGATTACAGCTCATTTGACGTCGAGGGCAAAAAGAGCTGGCTCAA 300
Db 80 ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln 99
QY 301 ATTTTCAACAATATCTCAGAAAGCAAAATTGGTACGAGAACTCTCTACTGTTGAATAC 360
Db 100 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
QY 361 TGCACCCCAACAGATATTTTCTCATGTTATTGAAGGTTATGAATCTCCAGAAATAGCT 420
Db 120 SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla 139
QY 421 CTAAATTTGTGAATAATGTTAAGAGAAATGCAATGCAATGAACCACTTGCAGAAATCAT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 159
QY 481 TTGTGGTCGGAACAGTTTATGATTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 160 LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATGCTATGTCACATTCACAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAATGTTACTTCA 660
Db 200 AspPheLeuGluGlnAsnTyrAspThrIlePheGluAspTyrGluLysLeuLeuGlnSer 219
QY 661 GAAATTTATGTGACAAAGACAGTCACTGAGCTTCTCGGTGAACCTACTAGATAGAGA 720
Db 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAATCTCACAATTATGACAAATATACATCAGTAAACCTCAGAACCTCAATTAATCATG 780
Db 240 HisAsnPheAlaIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACTGCTGCGAGACAAAGTCGCAACATCCAGTTTCAGGCTTTCACGTTTAAAGGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCCAATCTTAACAAGACGACGCCATCTAGACATCTCTCAAGAACCCAGGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 ARACTCATAGAGTCTCTCAGCAAGTTTCAGAACGACGAGCGAGTGTATGACAGTTCC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGACGAATTCCTCGGTCGATTTACGCTTAAACCGCGAGCGGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyrIleLysGlnIleArg 330
QY 1021 GATTGAAGACACAGCT 1038
Db 331 AspLeuLysLysThrAla 336
```

RESULT 6

```
US-10-025-730-4
; Sequence 4, Application US/10025730
; Publication No. US20030045466A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
```

```
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; PRIORITY FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE: -
; OTHER INFORMATION: g1794137
US-10-025-730-4

Alignment Scores:
Pred. No.: 2,07e-100 Length: 339
Score: 1117.00 Matches: 220
Percent Similarity: 84.95% Conservative: 51
Best Local Similarity: 68.97% Mismatches: 40
Query Match: 61.95% Indels: 8
DB: 14 Gaps: 5
```

US-10-089-688-1 (1-1053) x US-10-025-730-4 (1-339)

```
QY 1 ATCCGCTCCCGTTGGAGTCTCACAAATCTCCAGCACACATTGTGAAGATCTGAAG 60
Db 1 MetProLeu---PheGlyLysSerGlnLysSerProValGluLeuValLysSerLeuLys 19
QY 61 GAGAGCATGCTGCTTCTGGAAAAGCAGACATTTCTGATAAAAAGCAGAAAAGCGCTACA 120
Db 20 GluAlaIleAsnAlaLeuGlu-----AlaGlyAspArgLysValGluLysAlaGln 36
QY 121 GAAGAAGTTTCCAAAATCTGGTGGCATGAAAGAAATCTGTATGGCACAAATGAAAAA 180
Db 37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuHisGlySerSerAspAla 56
QY 181 GAGCCTCAGACAGAA---GCAGTAGCTCACTTGCCTCAGAACTCTATATAGTGGGCTC 237
Db 57 GluProProAlaAspTyrValValAlaGlnLeuSerGlnGluLeuTyrAsnSerAsnLeu 76
QY 238 CTTAGCACCTGCTAGTGTATTTACAGCTCATTGACTTTTCAGGCGCAAAAAGAGCTGCCT 297
Db 77 LeuLeuLeuLeuIleGlnAsnLeuHisArgIleAspPheGluGlyLysLysHisValAla 96
QY 298 CAAATTTTCAACAATATCTCAGAGACAAATTTGGTACGAGAACTCCTACTGTTGAATAC 357
Db 97 LeuIlePheAsnAsnLeuLeuArgGlnIleGlyThrArgSerProThrValGluTyr 116
QY 358 ATCTGCGACCCACAGAAATATTTGTTTCATGTTATTGAAAGGTTATGATCT-----CCA 411
Db 117 IleCysThrLysProGluIleLeuPheThrLeuMetAlaGlyTyrGluAspAlaHisPro 136
QY 412 GAAATAGCTCTAAATTTGCGAATAATGTTAAGAGAAATGCAATCAGACATGAACCACTTGCA 471
Db 137 GluIleAlaLeuAsnSerGlyThrMetLeuArgGluCysAlaArgTyrGluAlaLeuAla 156
QY 472 AAAATCATTTTGTGGTCCGAACAGTTTATGATTTTCTTCAGATATGTGCGAAATGTCAACA 531
Db 157 LysIleMetLeuHisSerAspGluPhePheLysPhePheArgTyrValGluValSerThr 176
QY 532 TTGACATAGCTTCAGATCCATTTGCCACATTTCCACAGGATTTACTTACAGACATTAATG 591
Db 177 PheAspIleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeu 196
QY 592 CTCAGTGCAGAAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAA---TATGAGAAG 648
Db 197 LeuCysAlaGluPheLeuAspAlaAsnTyrAspLysPhePheSerGlnHisTyrGlnArg 216
```

```
QY 649 TTACTTCATTCAAGAAATTATGTGACAAAAGACAGTCACTGAAGCTTCTCGTGAACCTA 708
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 217 LeuLeuAsnSerGluAsnValThrArgArgGlnSerLeuLysLeuGlyGluLeu 236
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 709 CTACTAGATAGACAACTTCACAAATTATGACAAAATACATCATGTAACCTGAGAACCTC 768
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 237 LeuLeuAspArgHisAsnPheThrValMetThrArgTyrIleSerGluProGluAsnLeu 256
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 769 AAATTATGATGAACCTGCTCGAGACAAAAGTCGCAACATCCAGTGTGAGGCCTTTCAC 828
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 257 LysLeuMetMetAsnMetLeuLysGluLysSerArgAsnIleGlnPheGluAlaPheHis 276
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 829 GTTTTAAAGTGTGTGTAGCAATCTCTAAAGACGAGCCCATCTAGACATCTCTCTC 888
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 277 ValPheLysValPheValAlaAsnProAsnLysProLysProLysProLysLeuLeu 296
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 889 AAGAACGAGCCCAACTCATAGAGTCTCTCAGCAAGTTTTCAGAACGACGACGGAT 945
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 297 ArgAsnGlnThrLysLeuValAspPheLeuThrAsnPheHisThrAspArgSerGlu 315

RESULT 7
US-10-025-730-5
; Sequence 5, Application US/10025730
; Publication No. US20030045468A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina A.
; TITLE OF INVENTION: CALCIUM BINDING PROTEIN
; FILE REFERENCE: PF-0635 US
; CURRENT APPLICATION NUMBER: US/10/025,730
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US/09/190,965
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE: -
; OTHER INFORMATION: g1255838
US-10-025-730-5

Alignment Scores:
Pred. No.: 2,84e-94 Length: 377
Score: 1054.50 Matches: 205
Percent Similarity: 77.44% Conservative: 49
Best Local Similarity: 62.50% Mismatches: 61
Query Match: 58.49% Indels: 13
DB: 14 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-025-730-5 (1-377)
QY 1 ATGCCGTTCCCTTTGGGAAGTCTCAAAATCTCAGCGAGACATGTGAGAAATCTGAAG 60
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1 MetProLeuLeuPheGlySerHisLysSerProAlaAspValLysThrLeuArg 20
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 61 GAGACATGCTGTTCTGCAAG-----CAAGAC 90
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 21 GluValLeuThrIleLeuAspLysLeuProProLysLeuAspLysAspGlyAsnIle 40
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 91 ATTCTGTATAAAAAGACAGAAAGGCTACAGAAAGTTTCCAAAATCTGGTGGCATG 150
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 41 GlnSerAspLysLysThrAspLysAlaLeuAspGluValSerLysAsnValAlaMetile 60
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 151 AAGAAATCTGTATGGCACAATGAAAAGACCTCAGACAGAA-----GCAGTA 201
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 61 LysSerPheIleTyrGlyAsnAspSerAlaGluProSerGluHisValValGlnVal 80
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 202 GCTCAACTGCTCAAGAACTCTATAATAGTGGCTCTCTTAGCACCCCTGGTAGCTGATTA 261
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 81 AlaGlnLeuAlaGlnGluValTyrAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100
QY 262 CAGCTCATTCAGCTTTGAGGGCAAAAAGACGCTGCGCTCAAATTTTCAACAATATCTCAGA 321
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 101 ProLysPheGluPheGluCysLysAspValGlyGlnIlePheAsnAsnLeuLeuArg 120
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 322 AGACAAAATGGTACAGAACTCTCTACTGTGTAATACATCTGCACCCCAACAAGATATTTTG 381
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyrLeuGlyAlaArgProGluIleLeu 140
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 382 TTTCATGTTTATGAAGGGTATGAATCTCCAGAAAATAGCTCTAAATGTGGAATATGTTA 441
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 141 IleGlnLeuValGlnGlyTyrSerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 442 AGAGATGATCAGACATGCAACCTTGCACAAAATCATTTTGTGTGCGGAAACAGATTTTAT 501
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTyrSerAspValPheTyr 180
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 502 GATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGTTCAGATGATGATTTGCCACA 561
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 181 ThrPhePheLeuTyrValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 562 TTCACAGATTTTACTTACAGACATAAATGTCTCAGTGCAGAAATTTTGGAAACAGCATTAT 621
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 201 PheLysGluLeuThrThrArgHisLysAlaIleAlaGlnPheLeuAspSerAsnTyr 220
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 622 GATAGATTTTTCAGTGAATATGAGAAAGTTACTTTCATTCAGAAAATATATGTGACAAAAGA 681
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 221 AspThrPheAlaGlnTyrGlnAsnLeuLeuAsnSerLysAsnTyrValThrArgArg 240
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 682 CAGTCACTGAAGCTTCTCGTGAACTACTACTAGATAGACACAACTTCACAAATATGACA 741
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 241 GlnSerLeuLysLeuLeuGlyGluLeuLeuAspArgHisAsnPheAsnThrMetThr 260
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 742 AAATACATCAGTAAACCTGAGAACCTCAATTAATGATGAACCTGCTGCGAGACAAAAGT 801
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 261 LysTyrIleSerAsnProAsnLeuArgLeuMetGluLeuLeuArgAspLysSer 280
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 802 CGCAACATCCAGTTGAGGCTTTCACGTTTAAAGTGTGTTAGCCAAATCCTTAACAAG 861
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 281 ArgAsnIleGlnTyrGluAlaPheHisValPheLysValPheValAlaAsnProAsnLys 300
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 862 ACGGACCCCATCTCCTAGACATCTCCTCAAGAACGAGCCCAAACTCATAGATTCCTCAGC 921
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 301 ProLysProIleSerAspIleLeuAsnArgAsnArgGluLysLeuValGluPheLeuSer 320
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 8
US-10-425-114-52177
; Sequence 52177, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 52177
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3912-007-H10_FLI pep
```


US-10-425-114-52177

Alignment Scores:

Pred. No.: 2.06e-64 Length: 446
Score: 750.00 Matches: 147
Percent Similarity: 66.98% Conservative: 64
Best Local Similarity: 46.67% Mismatches: 100
Query Match: 41.60% Indels: 4
DB: 15 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-425-114-52177 (1-446)

```
QY 13 TTTCGGAAGTCTCACAATCTCCAGCAGCATCTGGAAGATCTGAAGGAGCATGGCT 72
DB 115 PhelysSerLysProArgThrProValAspIleValArgGlnThrArgGluCysLeuVal 134
QY 73 GTTCTGGAAGACGACATTTCT- - - - -GATAAAAAGCAGAAAGGCTACAGAA 123
DB 135 TyrLeuAspLeuHisSerAspSerArgSerGlyAspAlaLysArgGluLysMetThr 154
QY 124 GAAGTTTCCAAAATCTGGTTCATGAAAGAAATCTGTATGGCACAATAATGAAAAAGAG 183
DB 155 GluLeuSerLysAenIleArgAspMetLysSerIleLeuTyrGlyAsnGlyGluSerGlu 174
QY 184 CCTCAGACAGACAGTAGCTCAACTGCTCAAGACTCTATATAGTGGGCTCCTTAGC 243
DB 175 ProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeuArg 194
QY 244 ACCCTGGTAGCTGATTTACAGCTCATTTGAGGCGCAAAAAGAGCTGGCTCAAAAT 303
DB 195 LeuLeuIleHisLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGlnVal 214
QY 304 TTCAACAATATCTCAGAGACAAATGGTACGAGAACTCCTACTGTTGAATACATCTGC 363
DB 215 ValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluTyrLeuGlu 234
QY 364 ACCCAACAGATATTTTGTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTTA 423
DB 235 SerAsnLysAspLeuLeuAspIleLeuIleLeuGlyTyrGluAsnMetAspIleAlaLeu 254
QY 424 AATGTGGAATATGTTAAAGAGATGATCAGACATCAACACCTTGCAGAAATCATTTTG 483
DB 255 HisTyrGlyAlaMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrValLeu 274
QY 484 TGGTCGGAACAGTTTATGATTTCTTCAGATATGCGAAATGTCAACATTTGACATAGCT 543
DB 275 GluSerGluHisMetLysLysPheAspTyrIleGlnLeuProAsnPheAspIleAla 294
QY 544 TCAGATCATTTGCCATCAATCAGGATTTACTTACAGACATATAATTTGCTCAGTGCAGAA 603
DB 295 SerAspAlaSerAlaThrPheLysGluLeuLeuThrArgHisLysAlaThrValAlaGlu 314
QY 604 TTTTGGACACGACATTATGATAGATTTTTCAGTGAATAT- - -GAGAAGTTTACTTCATCA 660
DB 315 PheLeuSerAsnAsnLysAspTyrPhePheGluGluPheAsnSerArgLeuLeuSerSer 334
QY 661 GAAATTTATGTGACAAAAGACAGTCTCACTGAAGCTTCTCGGTGAGAACTACTAGATAGA 720
DB 335 ThrAsnTyrIleThrLysArgGlnAlaIleLysLeuGlyAspMetLeuLeuAspArg 354
QY 721 CACAACTTCAATATGACAAAATATACATCAGTMAAAGCTCAGAACCTCAAAATTAATGATG 780
DB 355 SerAsnAlaAlaValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeuMet 374
QY 781 AACTGCTGAGACAAAAGTCGCAACATCCAGTTTTCAGGCTTTCACAGTTTTCAGGTTG 840
DB 375 AsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 394
QY 841 TTTGTAGCCCAATCTTAACAGAGCGCAGCCCATCTAGACATCTCTCCTCAAGAACCCAGGCC 900
DB 395 PheAlaAlaAsnLysAsnLysProGluValValAsnIleLeuValThrAsnArgSer 414
QY 901 AAACATCATAGAGTTCTTCAGCAAGTTTCAGAACGACGAGCGGAT 945
```

Db 415 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 429

RESULT 9

US-10-425-115-322057
; Sequence 322057, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 322057
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(337)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_5677C.1.pep
US-10-425-115-322057

Alignment Scores:

Pred. No.: 3.36e-64 Length: 337
Score: 747.50 Matches: 146
Percent Similarity: 66.46% Conservative: 64
Best Local Similarity: 46.20% Mismatches: 101
Query Match: 41.46% Indels: 5
DB: 17 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-425-115-322057 (1-337)

```
QY 13 TTTCGGAAGTCTCACAATCTCCAGCAGCATCTGGAAGATCTGAAGGAGCATGGCT 72
DB 5 PhelysSerLysProArgThrProValAspIleValArgGlnThrArgGluCysLeuVal 24
QY 73 GTTCTGGAAGACGAA- - - - -GACATTTCTGATAAAAAAGCAGAAAGGCTACA 120
DB 25 TyrLeuAsp***SerThrArgThrProGlyAla***AspAlaLysArgGluGluLysMet 44
QY 121 GAAGAAGTTTCCAAAATCTGGTTCATGAAAGAAATCTGTATGGCACAATAATGAAAA 180
DB 45 ThrGluLeuSerLysAenIleArgAspMetLysSerIleLeuTyrGlyAsnGlyGluSer 64
QY 181 GAGCCTCAGACAGACAGTAGCTCAACTGCTCAAGACTCTATATAGTGGGCTCCTT 240
DB 65 GluProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeu 84
QY 241 AGCACCTCTGTAGCTGATTTTACAGCTCATTTGAGGCGCAAAAAGACGTGGCTCAA 300
DB 85 ArgLeuLeuIleHisLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGln 104
QY 301 ATTTTCAACAATATCTCAGAGACAAATTTGGTACGAGAACTCCTACTGTTGAATACATC 360
DB 105 ValValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluTyrLeu 124
QY 361 TGACCCCAACAGATATTTTGTTCATGTTTATGAAAGGTTATGAAATCTCCAGAAATAGCT 420
DB 125 GluSerAsnLysAspLeuLeuAspIleLeuIleLeuGlyTyrGluAsnMetAspIleAla 144
QY 421 CTAAATTTGGGAATAATGTTAAGAGAAATGTCATCAGACCAATGAAACCATTTGCAAAATCAT 480
DB 145 LeuHisTyrGlyAlaMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrVal 164
QY 481 TTGTGTCGGAACAGGTTTTTATGATTTTTCAGATATGTCGAAATGTCAACATTTGACATA 540
```


Db 165 LeuGluSerGluHisMetLysLysPhePheAspTyrIleGlnLeuProAsnPheAspIle 184
Qy 541 GCTTCAGATCATTGTCACATTCAGGATTACTTACAGACATAAATTCGTCAGTGA 600
Db 185 AlaSerAspAlaSerAlaThrPheLysGluLeuLeuThrArgHisLysAlaThrValAla 204
Qy 601 GAATTTTGGACAGCAGATTATGATAGATTTTTCAGTGAATAT--GAGAGATTACTTCAT 657
Db 205 GluPheLeuSerAsnAsnTyrAspTTPhePheGluGluPheAsnSerArgLeuLeuSer 224
Qy 658 TCAGAAAATTTATGTGCAAAAAGACAGTCACTCAAGCTTCTCGTGAACCTACTACTAGAT 717
Db 225 SerThrAsnTyrIleThrLysArgGlnAlaIleLysLeuGluGlyAspMetLeuLeuAsp 244
Qy 718 AGACACAACTTCAATATGACAAATATACATCAGTAAACCTCAGAACCTCAATTAATG 777
Db 245 ArgSerAlaAlaValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeu 264
Qy 778 ATGAACTCTGGCAGACAAAGTCGCAACATCCAGTTTCAGGCTTTCACGTTTTTAAG 837
Db 265 MetAsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLys 284
Qy 838 GTGTTTCTAGCCATCTCAACAGCAGCCATCCTAGACATCCTCCTCAAGAACCGAT 897
Db 285 LeuPheAlaAlaAsnLysAsnLysProGluValValAsnIleLeuValThrAsnArg 304
Qy 898 GCCAACTCATAGAGTTCTCAGCAAGTTTCAGAACGACGAGGACGGAT 945
Db 305 SerLysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 320

RESULT 10

US-10-437-963-161075
; Sequence 161075, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161075
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET4530_60293C.1.rapb
US-10-437-963-161075

Alignment Scores:

Pred. No.: 3,76e-64 Length: 336
Score: 747.00 Matches: 146
Percent Similarity: 67.30% Conservative: 66
Best Local Similarity: 46.33% Mismatches: 99
Query Match: 41.43% Indels: 4
DB: 16 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-437-963-161075 (1-336)

Qy 13 TTGGAAGTCTCAAAATCTCCAGCAGCATTTGTGAAGATCTGAAGGAGAGCATGGCT 72
Db 5 PheLysSerLysProAlaArgProAlaValArgValValArgGlnThrArgGluLeuLeu 24
Qy 73 GTTCTGGAAAAGCAAGCATTTCT-----GATAAAAAGCAGAAAAGGCTACAGAA 123

Db 25 PheLeuAspLeuHisSerGlySerArgGlyGlyAspAlaLysArgGluGluLysMetAla 44
Qy 124 GAAGTTTCCAAAATCTGGTTCCTCCATGAAGAAAATTTCTGTATGGCACAATGAAAAGAG 183
Db 45 GluLeuSerLysAsnIleArgGluLeuLysSerIleLeuTyrGlyAsnGlyGluSerGlu 64
Qy 184 CCTCAGACAGAGCAGTACCTCAATTCGTCTCAAGAACTCTATAATAGTGGGCTCTTACG 243
Db 65 ProValThrGluAlaCysValGlnLeuThrGlnGluPhePheArgGluAsnThrLeuArg 84
Qy 244 ACCCTGGTAGCTGATTACAGCTCATTCAGCTTTGAGGCGCAAAAAGACGCTGGCTCAATT 303
Db 85 LeuLeuIleCysLeuProLysLeuAsnLeuGluThrArgLysAspAlaThrGlnVal 104
Qy 304 TTCACAAATATTCTCAGAACAGCAAAATGTTGTAAGAACTCTCTACTCTTGAATACATCTGC 363
Db 105 ValAlaAsnLeuGlnArgGlnValSerSerLysIleValAlaSerGluTyrLeuGlu 124
Qy 364 ACCCAACAGAAATATTTTGTTCATGTTTGAAGGGTATGAATCTCCGAAATAGTCTTA 423
Db 125 AlaAsnLysAspLeuLeuAspThrLeuIleSerGlyTyrGluAsnMetAspIleAlaLeu 144
Qy 424 AATTGGAATATGTTAGAGAAATGCATCAGACATGAACCTTCACAAAATCATTTTG 483
Db 145 HisTyrGlySerMetLeuArgGluCysIleArgHisGlnSerIleAlaArgTyrValLeu 164
Qy 484 TGTGTCGGAACAGTTTATGATTTCTCAGATATGCGAAATGTCAACATTTGACATAGCT 543
Db 165 GluSerAspHisMetLysLysPhePheAspTyrIleGlnLeuProAsnPheAspIleAla 184
Qy 544 TCAGATGTCATTTGCCACATTCAGAGATTTACTTACAGACATAAATTCGTCAGTGCAGAA 603
Db 185 SerAspAlaSerAlaThrPheLysGluLeuLeuThrArgHisLysAlaThrValAlaGlu 204
Qy 604 TTTTGGACAGCATTTATGATAGATTTTTCAGTGAATATGAG--AAGTTACTTTCATTCA 660
Db 205 PheLeuSerLysAsnTyrAspTTPhePheSerGluPheAsnThrArgLeuLeuSerSer 224
Qy 661 GAAATATTATGTGACAAAAGACAGCAGTCACTCAAGCTTCTCGTGAACTACTACTAGATAGA 720
Db 225 ThrAsnTyrIleThrLysArgGlnAlaIleLysPheLeuGlyAspMetLeuLeuAspArg 244
Qy 721 CACAACTTCAATATGACAAATATCATCATGATAAACCCTGAGAACCTCAATTAATGATG 780
Db 245 SerAsnSerThrValMetMetArgTyrValSerSerLysAspAsnLeuMetIleLeuMet 264
Qy 781 AACTGCTGCGACACAAAAGTCGCAACATCCAGTTTGGAGGCTTTTCACGTTTTTAAGGTG 840
Db 265 AsnLeuLeuArgAspSerSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 284
Qy 841 TTTGTAGCCCAATCTCAACAGACGACGCCATCTCTAGACATCTCTCAAGAACCCAGGCC 900
Db 285 PheAlaAlaAsnLysAsnLysProThrGluValValAsnIleLeuValThrAsnArgSer 304
Qy 901 AAACCTCATAGAGTTCTCAGCAAGTTTCAGAACGACGACGAGGACGGAT 945
Db 305 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 319

RESULT 11

US-10-425-114-54669
; Sequence 54669, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B

```

Qy 781 AACCTGCTGCGAGACAAAGTCGCAATCCAGTTTCAGGCGCTTCACGTTTATTAAGTG 840
Db 341 AsnLeuLeuArgAspSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 360
Qy 841 TTTCAGCCAAATCTCAACAGACGACGCCCATCTCAGACATCTCTCAAGAACCCAGGCC 900
Db 361 PheAlaAlaAsnLysAsnLysProGluValValAlaAsnIleLeuValThrAsnArgAsn 380
Qy 901 AAATCATAGATGTTCTCCAGCAAGTTTCAGAACGACAGGCGGAT 945
Db 381 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 395

RESULT 12
US-10-424-599-224185
; Sequence 224185, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224185
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_4446C.1.pep
US-10-424-599-224185

Alignment Scores:
Pred No.: 1,19e-61 Length: 339
Score: 721.50 Matches: 150
Percent Similarity: 67.31% Conservative: 58
Best Local Similarity: 48.54% Mismatches: 98
Query Match: 40.02% Indels: 3
DB: 15 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-424-599-224185 (1-339)
Qy 1 ATGCCGTTCCCGTTGGGAGTGCT-----CACAAATCTCCAGCAGACATTGTGAAGAT 54
Db 1 MetSerPheSerPhePheLysAlaLeuArgProLysThrProGlnGluValAlaLysSer 20
Qy 55 CTGAAGCAGAGCATGGCTGTTCTCGAAAGACAGACATTTCTGATAAAAAGCA--GAA 111
Db 21 IleLysGluSerLeuMetAlaLeuAspThrLysThrValValGluValLysAlaLeuGlu 40
Qy 112 AAGCCTACAGAAGAAGTTTCCAAAAATCTGGTTCGCCATGAAGAAATTTCTGTATGGCACA 171
Db 41 LysAlaLeuGluGluValGluLysAsnPheValThrMetArgThrMetLeuSerGlyAsp 60
Qy 172 AATGAAAAAGCCTCAGACAGACAGTAGCTCAACTGCTCAAGAACTCTATAATAGT 231
Db 61 GlyGluSerGluProAsnLeuAsnGlnValSerGlnLeuValGluGluIleCysLysGlu 80
Qy 232 GGGCTCCTTAGCACCCCTGGTAGCTGATTTTACAGCTCATTGAGCTTTGAGGGCAAAAAGAC 291
Db 81 AspValLeuThrLeuLeuIleHisLysLeuProIleLeuGlyTrpGluAlaArgLysAsp 100
Qy 292 GTGCTCAAAATTTTCAACAATATTTCTCAGAAGACAAATTTGGTAGCAGAACTCCTACTGTT 351
Db 101 LeuValHisCysTrpSerIleLeuLeuLysHisLysValGluThrAsnTyr-Tyr-CysVal 120
Qy 352 GAATACATCTGCACCCCAACAGATATATTTTGTTCATGTTATTATGAAGGGTATGAATCTCCA 411
Db 121 GluTrpIleGluGlnHisIleGluLeuLeuAspPheLeuValCysTyrAspAsnLys 140

```

```

Qy 781 AACCTGCTGCGAGACAAAGTCGCAATCCAGTTTCAGGCGCTTCACGTTTATTAAGTG 840
Db 341 AsnLeuLeuArgAspSerLysAsnIleGlnIleGluAlaPheHisValPheLysLeu 360
Qy 841 TTTCAGCCAAATCTCAACAGACGACGCCCATCTCAGACATCCTCTCAAGAACCCAGGCC 900
Db 361 PheAlaAlaAsnLysAsnLysProGluValValAsnIleLeuValThrAsnArgAsn 380
Qy 901 AAATCATAGATTTCTCCAGCAAGTTTCAGAACGACAGGCGGAT 945
Db 381 LysLeuLeuArgPhePheAlaGlyPheLysIleAspLysGluAsp 395

RESULT 12
US-10-424-599-224185
; Sequence 224185, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 39-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224185
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_44468C.1.pep
US-10-424-599-224185

Alignment Scores:
Pred No.: 1,19e-61 Length: 339
Score: 721.50 Matches: 150
Percent Similarity: 67.31% Conservative: 58
Best Local Similarity: 48.54% Mismatches: 98
Query Match: 40.02% Indels: 3
DB: 15 Gaps: 2

US-10-089-688-1 (1-1053) x US-10-424-599-224185 (1-339)
Qy 1 ATGCCGTTCCCGTTGGGAGTGCT-----CACAAATCTCCAGCAGACATTGTGAAGAT 54
Db 1 MetSerPheSerPhePheLysAlaLeuArgProLysThrProGlnGluValAlaLysSer 20
Qy 55 CTGAAGCAGAGCATGGCTGTTCTCGAAAGACAGACATTTCTGATAAAAAGCA--GAA 111
Db 21 IleLysGluSerLeuMetAlaLeuAspThrLysThrValValGluValLysAlaLeuGlu 40
Qy 112 AAGCCTACAGAAGAAGTTTCCAAAAATCTGGTTCGCCATGAAGAAATTTCTGTATGGCACA 171
Db 41 LysAlaLeuGluGluValGluLysAsnPheValThrMetArgThrMetLeuSerGlyAsp 60
Qy 172 AATGAAAAAGCCTCAGACAGACAGTAGCTCAACTGCTCAAGAACTCTATAATAGT 231
Db 61 GlyGluSerGluProAsnLeuAsnGlnValSerGlnLeuValGluGluIleCysLysGlu 80
Qy 232 GGGCTCCTTAGCACCCCTGGTAGCTGATTTTACAGCTCATTGAGCTTTGAGGGCAAAAAGAC 291
Db 81 AspValLeuThrLeuLeuIleHisLysLeuProIleLeuGlyTrpGluAlaArgLysAsp 100
Qy 292 GTGCTCAAAATTTTCAACAATATTTCTCAGAAGACAAATTTGGTAGAGAACTCCTACTGTT 351
Db 101 LeuValHisCysTrpSerIleLeuLeuLysHisLysValGluThrAsnTyr-Tyr-CysVal 120
Qy 352 GAATACATCTGCACCCCAACAGATATATTTTGTTCATGTTATTATGAAGGGTATGAATCTCCA 411
Db 121 GluTrpIleGluGlnHisIleGluLeuLeuAspPheLeuValCysTyrAspAsnLys 140

```


Alignment Scores:			
Pred. No.:	1.29e-53	Length:	337
Score:	639.50	Matches:	135
Percent Similarity:	62.28%	Conservative:	73
Best Local Similarity:	40.42%	Mismatches:	109
Query Match:	35.47%	Indels:	17
DB:	16	Gaps:	2
US-10-089-688-1 (1-1053) x US-10-437-963-183779 (1-337)			
QY	31	TCTCCAGCAGACATTGTGAAGAATCTCAAGCAGACGATGCTGTTCTGGAAAGCAAGAC	90
DB	15	SerProGlnGluLeuValArgSerIleGlyGluSerLeuAlaLeuasp-----	31
QY	91	ATTTCTGATAAAAGCAGAAAAGGCTACAGAGAAGTTTCCAAAATCTGGTTGCCATG	150
DB	32	-----ThrArgThrGlyAlaLysAlaLeuGluaspValGluLysAsnValSerThrLeu	49
QY	151	AAAGAAATTCGTATGCGCACAAATGAAAAGAGCGCTCAGACAGACGATGCTCAACTT	210
DB	50	ArgGlnThrLeuSerGlyAspGlyGluValGluProAsnGlnGluGlnValLeuGlnIle	69
QY	211	GCTCAAGAACCTCTATAATAGTGGGCTCCTTAGCACCCCTGGTAGCTGATTTTACAGCTCATT	270
DB	70	AlaLeuGluIleCysLysGluaspValLeuSerLeuPheValGlnAsnMetProSerLeu	89
QY	271	GACTTTGAGGCAAAAAGACGTGGCTCAAATTTTCAACAATATTTCTCAGACAGACAATT	330
DB	90	GlyTrpGluGlyArgLysAspLeuAlaHisCysTrpSerIleLeuLeuArgGlnLysVal	109

Search completed: November 10, 2004, 21:14:19
Job time : 184.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:34:23 : Search time 47.5 Seconds
(without alignments)
4265.947 Million cell updates/sec

Title: US-10-089-688-1
Perfect score: 1803
Sequence: 1 atgcgcgtccgcttggaa.....cagctcagaagaagcttaa 1053

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model
-Q=/cgn2 1/USPTO.spool p/US10089688/runat 10112004 152247 4861/app.query.fasta_1.1223
-DB=PIR.79 -QFWT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -PART=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THE SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10089688 @CGN 1.1_63 @runat 10112004 152247 4861 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR.79.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1614.5	89.5	341	2	157997
2	1054.5	58.5	377	2	T16651
3	1006	55.8	338	2	T27129
4	766	42.5	329	2	T50117
5	666.5	37.0	305	2	G71441
6	614	34.1	348	2	B84448
7	501	27.8	399	2	S34681
8	137.5	7.6	339	2	T33477
9	122	6.8	602	2	T03057
10	117	6.5	365	2	D97848
11	115	6.4	2819	2	A90551
12	113.5	6.3	1939	2	T18372
13	113	6.3	2166	2	G70163
14	112.5	6.2	677	2	H64574

15	111	6.2	4981	2	T18489
16	110.5	6.1	4131	2	T21085
17	108.5	6.0	719	2	A81358
18	108	6.0	2401	2	T28676
19	107.5	6.0	1050	2	A89769
20	106.5	5.9	348	2	D70195
21	106	5.9	297	2	T28188
22	106	5.9	824	2	T20477
23	105.5	5.9	1619	2	T18499
24	105	5.8	809	1	JQ0032
25	105	5.8	871	2	S68482
26	104.5	5.8	617	2	D96978
27	104.5	5.8	1156	2	B70356
28	104	5.8	299	1	F70350
29	104	5.8	592	2	T29402
30	103.5	5.7	415	2	A41253
31	103.5	5.7	1189	2	A54817
32	103.5	5.7	1642	2	T08880
33	103.5	5.7	2473	1	S38040
34	103	5.7	256	2	B81394
35	103	5.7	355	2	C64644
36	103	5.7	799	2	T00331
37	103	5.7	959	2	T00246
38	103	5.7	2104	2	T38774
39	102.5	5.7	472	2	T19375
40	102.5	5.7	600	2	T09676
41	102.5	5.7	1300	2	T18364
42	102.5	5.7	1666	2	T38393
43	102	5.7	717	2	T27066
44	102	5.7	717	2	T27067
45	102	5.7	878	2	T23537

ALIGNMENTS

RESULT 1

I57997

hypothetical calcium-binding protein - mouse

C;Species: Mus sp. (mouse)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 19-May-2000

C;Accession: I57997

R;Miyamoto, H.; Matsushiro, A.; Nozaki, M.

Mol. Reprod. Dev. 34, 1-7, 1993

A;Title: Molecular cloning of a novel mRNA sequence expressed in cleavage stage mouse

A;Reference number: I57997; MUID:93119656; PMID:8418809

A;Accession: I57997

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-341 <RES>

A;Cross-references: GB:S51858; NID:G262933; PIDN:AAB24801.1; PID:G262934

C;Superfamily: Saccharomyces hypothetical protein YK1189W

C;Keywords: calcium binding

Alignment Scores:

Pred. No.: 1.1e-116 Length: 341

Score: 1614.50 Matches: 325

Percent Similarity: 93.43% Conservative: 2

Best Local Similarity: 92.86% Mismatches: 14

Query Match: 89.55% Indels: 9

DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x I57997 (1-341)

QY 1 ATCCCGTCCCGTTCCGAGTCTCAAAATCTCCAGCAGACATTGTGAAGATCTGAAG 60

Db 1 MetProPhProPhGlySerHisLysSerProAlaAspIleVallyAsnLeuLys 20

QY 61 GAGCAGATGCTCTTCTGGAAGACAGACATTCTGATAAAAAAGCAGAAAAAGGCTACA 120

Db 21 GluSerMetAlaValleuGluLysGlnAspIleSerAspLysAlaGluLysAlaThr 40

QY 121 GAAGAAGTTTCCAAAATCTGGTTGCATGAAGAAGAAATCTGTATGCGACAAATGAAAA 180

|||||

Db 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuTyrGlyThrAsnGluLys 60
Qy 181 GAGCCTCAGACAGAGCAGTAGCTCAACTGCTCAAGAACTCTATATAGTAGTGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLeuTyrAsnSerGlyLeuLeu 80
Qy 241 AGCACCTGGTAGCTGATTTACAGCTCATTTGAGGGCAAAAGACGTCGCTCAA 300
Db 81 GlyThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
Qy 301 ATTTTCAACATATCTCAGAGACAAATTTGCTAGAGAACTCTACTGTTGATATCATC 360
Db 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
Qy 361 TGCACCAACAGATATTTGTTTCATCTTATGAAAGGATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
Qy 421 CTAATTTGTGAATAATGTTAAGAGAATGCATCAGACATGAACACCTTGCAGAAATCAT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
Qy 481 TTGTGGTCGGAACAGTTTTATGATTTCTCAGATATGTCGAATGTCGAATTTGACATA 540
Db 161 LeuTrpSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
Qy 541 GCTTCAGATGCTATGTCACATTCACAGATTTACTTACAGACATAAATGCTCAGTGCA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
Qy 601 GAATTTTGGCAACAGCATTTATGATGATTTTTCAGTGAATATGAGAATGTTACTTCATCA 660
Db 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAATATGATGACAAAGACAGTCACTAGAGCTTCTCGTGAACTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
Qy 721 CACAACTTCAATATGACAAATATACATCATGATAAAGCTGAGAACCTCAATTAATCATG 780
Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACCTGCTCGAGACAAAGTCGCAACATCAGTTTGAAGGCTTTCAGGCTTTTAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCCAATCTTAAACAGACGAGCCCATCTAGACATCTCTCAAGAACGAGGCC 900
Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAACCTATAGAGTTCCTCAGCAAGTTTCAGAACGACGACGAGGATTTGTATGAGCAGTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
Qy 961 GTACCGCAGCAATTCCTCGGCTCGATTTACGCGTTAAACCGCGAGCGCTGGGATCAGG 1020
Db 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
Qy 1021 GATTGTAGAGACACGCTCAGCAGAGAGCT 1050
Db 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341

RESULT 2

Tl6651

Hypothetical protein R02E12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: Tl6651

R:Leimbach, D.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid R02E12.

A:Reference number: Zl8554

A:Accession: Tl6651

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-377 <LEI>

A:Cross-references: UNIPROT:Q21643; EMBL:U53337; NID:g1255833; PID:g1255838; PIDN:AAA961

A:Experimental source: strain Bristol N2; clone R02E12

C:Genetics:

A:Gene: CESP:R02E12.2

A:Map position: X

A:Introns: 37/3; 146/2; 225/1; 315/3

C:Superfamily: Saccharomyces hypothetical protein YKL189w

Alignment Scores:

Pred. No.:	1,6e-73	Length:	377
Score:	1054.50	Matches:	205
Percent Similarity:	77.44%	Conservative:	49
Best Local Similarity:	62.50%	Mismatches:	61
Query Match:	58.49%	Indels:	13
DB:	2	Gaps:	2

US-10-089-688-1 (1-1053) x Tl6651 (1-377)

Qy 1 ATCCCTTCCCGTTTCGGAAGTCTCAAAATCTCCAGACAGACATTTGTGAAGATCTGAAG 60

Db 1 MetProLeuLeuPheGlyLysSerHisLysSerProAlaAspValValLysThrLeuArg 20

Qy 61 GAGACATGCGCTCTTCTGAAAG-----CAGAC 90

Db 21 GluValLeuThrIleLeuAspLysLeuProProLysLeuAspLysAspGlyAsnIle 40

Qy 91 ATTTCGATATAAAGCAGAAAGGCTCAGAGAGAGTTTCCAAAATCTGGTTGCCATG 150

Db 41 GlnSerAspLysLysTyrAspLysAlaLeuAspLysValSerLysAsnValAlaMetIle 60

Qy 151 AAAGAAATTCGTATGCGCAAAATGAAAGAGAGCCCTCAGACAGAA-----GCAGTA 201

Db 61 LysSerPheIleTyrGlyAsnAspSerAlaGluProSerSerGluHisValValGlnVal 80

Qy 202 GCTCAACTGCTCAAGAACTCTATATAGTGGGCTCTTAGCACCTCTAGCTGAGCTGATTA 261

Db 81 AlaGlnLeuAlaGlnGluValTyrAsnAlaAsnIleLeuProMetLeuIleLysMetLeu 100

Qy 262 CAGCTCATTCACCTTGGCGCAAAAGAGCGCTCAAAATTTTCAACAATATTTCTCAGA 321

Db 101 ProLysPheGluPheGluCysLysLysAspValGlyGlnIlePheAsnAsnLeuLeuArg 120

Qy 322 AGACAAATTCGTACGGAACCTCTACTGTTGAATACATCTGACCAACCAAGATATTTG 381

Db 121 ArgGlnIleGlyThrArgSerProThrValGluTyrLeuGlyAlaArgProGluIleLeu 140

Qy 382 TTCATGTTATTGAAGGATGAAATCTCCAGAAATAGCTCTAAATTTGTGAATAATGTTA 441

Db 141 IleGlnLeuValGlnGlyTyrSerValProAspIleAlaLeuThrCysGlyLeuMetLeu 160

Qy 442 AGAAGATGCTCAGACATGAACCACTTGCAAAATCATTTTGTGTCGGAACAGCTTTTAT 501

Db 161 ArgGluSerIleArgHisAspHisLeuAlaLysIleIleLeuTyrSerAspValPheTyr 180

Qy 502 GATTTCCTCAGATATGCGAAATGTCACCAATTTGACATAGCTTCAGATGCTATTCACCA 561

Db 181 ThrPhePheLeuTyrValGlnSerGluValPheAspIleSerSerAspAlaPheSerThr 200

Qy 562 TTCAGGATTTACTTCAAGACATAAATTCGTCAGTCAGCAATTTTGGACACAGCATAT 621

Db 201 PheLysGluLeuThrThrArgHisLysAlaIleAlaGluPheLeuAspSerAsnTyr 220

Qy 622 GATAGATTTTTCAGTGAATGAGAGCTTACTTTCATTTCAGAAAATTTATGTGACAAAAAGA 681

Db 221 AspThrPhePheAlaGlnTyrGlnAsnLeuLeuAsnSerLysAsnTyrValThrArgArg 240

Qy 682 CAGTCACTGAGCTTCTCGGTGAACACTACTAGATAGACACAACTTCACAAATTTATGACA 741

Db 241 GlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArgHisAsnPheAsnThrMetThr 260

QY 742 AATACATCATGTAACCTCAAAATTAATGATGAACCTGCTCGAGACAAAGT 801
|||||
Db 261 LysTyrIleSerAsnProAspAsnLeuArgLeuMetGluLeuLeuArgAspLysSer 280
|||||
QY 802 CGCAACATCCAGTTGAGGCGCTTTACGCTTTTAAAGGTGTTGTAGCCAAATCCCTAAACAAG 861
|||||
Db 281 ArgAsnIleGlnTyrGluAlaPheHisValPheLysValPheValAlaAsnProAsnLys 300
|||||
QY 862 ACGGAGCCATCCATAGCATCTCTCTCAAGAACAGCCCAAACTCATAGAGTTCTCTCAGC 921
|||||
Db 301 ProLysProIleSerAspIleLeuAsnArgAsnArgGluLysLeuValGluPheLeuSer 320
|||||
QY 922 AAGTTTCAGACGACGACGACGGAT 945
|||||
Db 321 GluPheHisAsnAspArgThrAsp 328
|||||
RESULT 3
T27129
hypothetical protein Y53C12A.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27129
R:Kershaw, J.; Lennard, N.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z20315
A:Accession: T27129
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-338 <WIL>
A:Cross-references: UNIPROT:O18211; EMBL:Z99277; PIDN:CAB16486.1; GSPDB:GN00020; CESP:Y5
A:Experimental source: clone Y53C12A
C:Genetics:
A:Gene: CESP:Y53C12A.4
A:Map position: 2
A:Introns: 29/3; 103/3; 136/2; 215/1; 282/3
C:Superfamily: Saccharomyces hypothetical protein YKL189w
Alignment Scores:
Pred. No.: 8.67e-70 Length: 338
Score: 1006.00 Matches: 188
Percent Similarity: 77.96% Conservative: 56
Best Local Similarity: 60.06% Mismatches: 67
Query Match: 55.80% Indels: 2
Gaps: 1
US-10-089-688-1 (1-1053) x T27129 (1-338)
QY 13 TTTGGGAAGTCTCACAATCTCCAGCAGACATCTGGAAGATCTGAGGAGCATGGCT 72
|||||
Db 6 PheGlyLysAlaAspLysThrProAlaAspValValLysAsnLeuArgAspAlaLeuLeu 25
|||||
QY 73 GTTCTGAAAAGCAA-----GACATTTCTGATAAAAGCAGAAAAGCGCTACAGAGAA 126
|||||
Db 26 ValIleAspArgHisGlyThrAsnThrSerGluArgLysValGluLysAlaIleGluGlu 45
|||||
QY 127 GTTTCGAAAATCTGGTGCATGAAAGAAATCTGTATGGCACAATGAAAAGAGCCT 186
|||||
Db 46 ThrAlaLysMetLeuAlaLeuAlaLysThrPheIleTyrGlySerAspAlaAsnGluPro 65
|||||
QY 187 CAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATATAGTGGGCTCTTAGCACC 246
|||||
Db 66 AsnAsnGluGlnValThrGlnLeuAlaIleGluValTyrAsnAlaAsnValLeuProMet 85
|||||
QY 247 CTGGTAGCTGATTACAGCTCATTGCTTTGAGGCGCAAAAAGACGCTGGCTCAAAATTTTC 306
|||||
Db 86 LeuIleLysHisLeuHisLysPheGluPheGluCysLysLysAspValAlaSerValPhe 105
|||||
QY 307 AACATATTCTCAGACAGAAATTTGGTACGAGAACTCTCTACTGTGTGAATCATCTGCACC 366
|||||
Db 106 AsnAsnLeuLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrLeuAla 125
|||||
QY 367 CAACAGATATTTGTTTCATGTTATTGAAAGGCTATGAATCTCCAGAAATAGCTCTAAT 426
|||||

Db 126 ArgProGluIleLeuIleThrLeuLeuLeuGlyTyrGluGlnProAspIleAlaLeuThr 145
QY 427 TGTGGAATAATGTTAAGAGAAATGCATCAGATCAACCACTTGCAGAAATCATTTTGTGG 486
|||||
Db 146 CysGlySerMetLeuArgGluAlaValArgHisGluHisLeuAlaArgIleValLeuTyr 165
|||||
QY 487 TCGGAACAGTTTATGATTTCTTCAGATATGTGGAATGTCAACATTTGACATAGCTTCA 546
|||||
Db 166 SerGluTyrPheGlnArgPheValPheValGlnSerAspValPheAspIleAlaThr 185
|||||
QY 547 GATGCATTTCCCATTCACAGGATTTACTTTACAGACATAAATTTGCTCAGTGCAGAAATTT 606
|||||
Db 186 AspAlaPheSerThrPheLysAspLeuMetThrLysHisLysAsnMetCysAlaIleGluTyr 205
|||||
QY 607 TTGGAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTTACTTCAATTCAGAAAAT 666
|||||
Db 206 LeuAspAsnAsnTyrAspArgPheGlyGlnTyrSerAlaLeuThrAsnSerGluAsn 225
|||||
QY 667 TATGTGACAAAAAGACAGCTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACAAAC 726
|||||
Db 226 TyrValThrArgArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArgHisAsn 245
|||||
QY 727 TTCACAATATGACAAAAATACATCATGATAACCTGAGAACCTCAAAATTAATGATGAACCTG 786
|||||
Db 246 PheSerThrMetAsnLysTyrIleThrSerProGluAsnLeuLysThrValMetGluLeu 265
|||||
QY 787 CTCGAGACAAAAAGTCGCAACATCCAGTTTGTAGGCTTTTCACGTTTTTAAAGTGTGTGTA 846
|||||
Db 266 LeuArgAspLysArgArgAsnIleGlnTyrGluAlaPheHisValPheLysIlePheVal 285
|||||
QY 847 GCCAATCTTCAAGAGCGAGCCCTAGACATCTCTCCTCAAGAACCGAGGCAAACTC 906
|||||
Db 286 AlaAsnProAsnLysProArgProIleThrAspIleLeuThrArgAsnArgAspLysLeu 305
|||||
QY 907 ATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGACGACGAT 945
|||||
Db 306 ValGluPheLeuThrAlaPheHisAsnAspArgThrAsn 318
|||||
RESULT 4
T50117
mo25 homolog [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C:Accession: T50117
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25039
A:Accession: T50117
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-329 <SEE>
A:Cross-references: UNIPROT:Q9P708; EMBL:AL157734; PIDN:CAB75774.1; GSPDB:GN00066; SPDB:8
A:Experimental source: strain 972h(-); cosmid c1834
C:Genetics:
A:Gene: SPDB:SPAC1834.06c
A:Map position: 1
A:Introns: 34/3; 185/3
C:Superfamily: Saccharomyces hypothetical protein YKL189w
Alignment Scores:
Pred. No.: 2.72e-51 Length: 329
Score: 766.00 Matches: 159
Percent Similarity: 68.37% Conservative: 55
Best Local Similarity: 50.80% Mismatches: 95
Query Match: 42.48% Indels: 4
Gaps: 2
US-10-089-688-1 (1-1053) x T50117 (1-329)
QY 1 ATGCCGTTCCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGTGAAGAAATCTGAAG 60
|||||
Db 1 MetSerPheLeuPheAsnLysArgProLysSerThrGlnAspValValArgCysLeuCys 20
|||||

61 GAGAGCATGGCTCTTCTGGAAAGCAGACATTTCTGATAAAAGCAGAAAGCGCTACA 120
: : : : :
21 AspAsnLeuProLysLeuGluLeuAsn-----AsnAspLysLys-----LysSerPhe 36
: : : : :
121 GAGAGATTTCCAAAATCTGGTTGCCATGAAAGAAATTTCTGATGGCACAATGAAAAA 180
: : : : :
37 GluGluValSerLysCysLeuGlnAsnLeuGluPheGluSerLysLysAspThrGlyLeu 56
: : : : :
181 GAGCCTCAGACAGACAGTAGCTCAACTTTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
: : : : :
57 GluProAspAlaAspLeuValSerAspLeuSerPheGlnLeuTyrGlnSerAsnLeuPro 76
: : : : :
241 AGCACCTGTGTAGCTGATTTAGCTCATTGCTTTCAGGGCAAAAAGACGTCGCTCAA 300
: : : : :
77 PheLeuLeuValArgTyrLeuProLysLeuGluPheGluSerLysLysAspThrGlyLeu 96
: : : : :
301 ATTTTCAACATATTTCTCAAGACAAATGTTGGTACGAGAACTCTACTGTTGAATACATC 360
: : : : :
97 IlePheSerAlaLeuLeuArgGHisValAlaSerArgTyrProThrValAspTyrMet 116
: : : : :
361 TGCACCCCAACAGAAATATTTGTTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCT 420
: : : : :
117 LeuAlaHisProGlnIlePheProValLeuValSerTyrTyrArgTyrGlnGluValAla 136
: : : : :
421 CTRAAATGTGGATATGTTAAGAGATGTCATCAGACATCAACCTTGCAGAAATCAATT 480
: : : : :
137 PheThrAlaGlySerIleLeuArgGluCysSerArgHisGluAlaLeuAsnGluValLeu 156
: : : : :
481 TTGTGGTCGGAACAGTATTTGATTTCTTCAGATATGTCGAAATGTCAACATTTGCACATA 540
: : : : :
157 LeuAsnSerArgAspPheThrPhePheSerLeuIleGlnAlaSerSerPheAspMet 176
: : : : :
541 GCTTCAGATGCTATGTCACATTCAGGATTTACTTCAAGACATATAAATGCTCAGTGCA 600
: : : : :
177 AlaSerAspAlaPheSerThrPheLysSerIleLeuLeuAsnHisLysSerGlnValAla 196
: : : : :
601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAGTACTTCAATCA 660
: : : : :
197 GluPheIleSerTyrHisPheAspGluPhePheLysGlnTyrThrValLeuLeuLysSer 216
: : : : :
661 GAAATATATGACAAAAGACAGTCACTGAAGCTTCTCGGTGAACACTACTACTAGATAGA 720
: : : : :
217 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGluGlyGluIleLeuLeuAsnArg 236
: : : : :
721 CACAACTTCACATATATGACAAAATACATCAGTAAACCTCGAGAACCTCAATTAATCATG 780
: : : : :
237 AlaAsnArgSerValMetThrArgTyrIleSerSerAlaGluAsnLeuLysLeuMetMet 256
: : : : :
781 AACCTGCTGCAGACAAAAGTCGCACATCCAGTTCAGGCTTTCAGGCTTTTAAAGTG 840
: : : : :
257 IleLeuLeuArgAspLysSerLysAsnIleGlnPheGluAlaPheHisValPheLysLeu 276
: : : : :
841 TTTGTAGCAATCTCAACAGCAGCGCCCATCTAGACATCCCTCCTCAAGAACCCAGGCC 900
: : : : :
277 PheValAlaAsnProGluLysSerGluGluValIleGluIleLeuArgArgAsnLysSer 296
: : : : :
901 AAATCATATAGTTCCTCAGCAAGTTTCAGAACGACGAGG 939
: : : : :
297 LysLeuIleSerTyrLeuSerAlaPheHisThrAspArg 309
: : : : :
RESULT 5
G71441
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
A:Variety: Columbia
C:Accession: G71441
R:Sevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weizenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erthof, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And

C.: Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: G71441
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-305 <BEV>
A:Cross-references: GB:Z97343; NID:g2245073; PID:e327051; PID:g2245086
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Saccharomyces hypothetical protein YKL189w

Alignment Scores:
Pred. No.: 1-26e-43 Length: 305
Score: 666.50 Matches: 136
Percent Similarity: 65.68% Conservative: 63
Best Local Similarity: 44.88% Mismatches: 93
Query Match: 36.97% Indels: 11
DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x G71441 (1-305)
QY 124 GAAGTTTCCAAAATCTGTTGCCATGAAAGAAATTTCTGATGGCACAATGAAAAAGAG 183
Db 8 GluLeuSerLysSerIleArgAspLeuLysLeuIleLeuTyrGlyAsnSerGluAlaGlu 27
QY 184 CCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGGCTCCTTAGC 243
Db 28 ProValAlaGluAlaCysAlaGlnLeuThrGlnGluPhePheLysAlaAspThrLeuArg 47
QY 244 ACCCTGTGTAGCTGATTTACAGCTCATTTGAGGGCAAAAAGACGTCGCTCAAAAT 303
Db 48 ArgLeuLeuThrSerLeuProAsnLeuAsnLeuGluAlaArgLysAspAlaThrGlnVal 67
QY 304 TTCACATATTTCTCAGACAGAAATGTTGGTACGAGAACTCTACTGTTGAATACATCTGC 363
Db 68 ValAlaAsnLeuGlnArgGlnGlnValAsnSerArgLeuIleAlaAlaAspTyrLeuGlu 87
QY 364 ACCCAACAGAAATATTTGTTTCATGTTATGAAAGGGTATGAATCTCCAGAAATAGCTCTA 423
Db 88 SerAsnIleAspLeuMetAspPheLeuValAspGlyPheGluAsnThrAspMetAlaLeu 107
QY 424 AATGTGGAAATAATGTTAAGAGAAATGTCATCAGACATGACCACTTGCAGAAAATCAATTTG 483
Db 108 HisTyrGlyThrMetPheArgGluCysIleArgHisGlnIleValAlaLysTyrValLeu 127
QY 484 TGGTCGGAACAGTATTTATGATTTCTCAGATATGTCGAAATGTCAACATTTGCATAGCT 543
Db 128 AspSerGluHisValLysLysPhePheTyrTyrIleGlnLeuProAsnPheAspIleAla 147
QY 544 TCAGATGCTATTTGCCACATTTCAAGGATTTACTTACAGACATATAATTTGCTCAGTCAGAA 603
Db 148 AlaAspAlaAlaAlaThrPheLysGluLeuLeuThrArgHisLysSerThrValAlaGlu 167
QY 604 TTTTGGACAGCATTAATAGATATTTTTCAGTGAATAT---GAGAGTTACTTCAATCA 660
Db 168 PheLeuIleLysAsnGluAspTyrPhePheAlaAspTyrAsnSerLysLeuLeuGluSer 187
QY 661 GAAATATGTCGACAAAAGACAGTCACTGAAGCTTCTCGTGAACACTACTACTAGATAGA 720
Db 188 ThrAsnTyrIleThrArgArgGlnAlaIleLysLeuGlyAspIleLeuLeuAspArg 207
QY 721 CACAACTTCACAAATATGACAAAATATACATCAGTAAACCTGAGAACTCTCAAAATTAATGATG 780
Db 208 SerAsnSerAlaValMetThrLysTyrValSerSerMetAspAsnLeuArgIleLeuMet 227
QY 781 AACCTGCTGCAGACAAAAGTCGCAACATCCAGTTTGGGCTTTCAGGCTTTTAAAGTG 840
Db 228 AsnLeuLeuArgGluSerSerLysThrIleGlnIleGluAlaPheHisValPheLysLeu 247
QY 841 TTTTGTAGCAATCTCAACAGCAGCGCCCATCTAGACATCCCTCCTCAAGAACCCAGGCC 900
Db 248 PheValAlaAsnGlnAsnLysProSerAspIleAlaAsnIleLeuValAlaAsnArgAsn 267

	Qy	418	GCTCTAAATTTGGGATAAATGTTAAGAGAATGCATCAGACATGAACCACTTTCGAAAAATC	477
	Db	145	AlaLeuHisIstyrThrGlyMetLeuLysGluCysValArgHisGlnValValAlaLysIstyr	164
	Qy	478	ATTMTTGTTGGTCGGAACACGTTTTATGATTTCTTCAGATATGTCGAATGTCAACATTTGAC	537
	Db	165	IleLeuGluSerLysAsnLeuGluLysPhePheAspTyrValGlnLeuProTyrPheAsp	184
	Qy	538	ATAGCTTCAGATGCATTTGCCACCATTTCAAGGATTTACTTTACAAGACATAAAATTCCTCAGT	597
	Db	185	VallalThrAspAlaSerLysIlePheArgGluLeuLeuThrArgHisLysSerThrVal	204
	Qy	598	GCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATAG-- --AAGTTACTT	654
	Db	205	AlaGluTyrLeuAlaLysAsnTyrGluTrpPhePheAlaGluTyrAsnThrLysLeuLeu	224
	Qy	655	CATTCCAGAAATTTATGTGACAAAAAGACAGTCATCGAAGCTTCTCGTGAACACTACTACTA	714
	Db	225	GluLysGlySerTyrPheThrLysArgGlnAlaSerLysLeuLeuGlyAspValLeuMet	244
	Qy	715	GATGACACACATCTCACAAATTTATGACAAAATACATCATCAACCTTGAGAACCTCAAATTA	774
	Db	245	AspArgSerAsnSerglyValMetVallyIstyrValSerSerLeuAsnLeuArgIle	264
	Qy	775	ATGATGAACCTGCTCGCAGACAAAAGTCGCAACATCCAGTTTGAGCGCTTTCACCGTTTTT	834
	Db	265	MetMetAsnLeuLeuArgGluProThrLysAsnIleGlnLeuGluAlaPheHisIlePhe	284
	Qy	835	AAGGTGTTGTAGCCAATFCCTTAACAAGACGCGCCCATCTCTAGACATCTCTCTCAAGAAC	894
	Db	285	LysLeuPheValAlaAsnGluAsnLysProGluAspIleValAlaIleLeuValAlaAsn	304
	Qy	895	CAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTTCAGAACGACGACGCGAT	945
	Db	305	ArgThrIstyrIleLeuGluPheAlaAspLeuLysProGluLysGluAsp	321

Db 238 PheAsnSerAlaProIleHisTyrGlnGlySerProAsnGluLysGluAspAlaGln 257
Qy 994 GTTAAACCGCGAGCGCTGGGATCAGG-----GATTGGAAGACACGAGCTCGACAGAA 1047
Db 258 TyrAlaArgMetAlaTyrLysLeuAsnTrpAspMetGlnArgProPheThrGlnGlu 277
RESULT 9
T03057
hypothetical protein 032R - Chilo iridescent virus
C:Species: Chilo iridescent virus
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03057
R:Bahr, U.; Tidona, C.A.; Barai, G.
Virus Genes 15, 235-245, 1997
A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101
A:Reference number: Z14834; MUID:98141693; PMID:9482589
A:Accession: T03057
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-602 <BAH>
A:Cross-references: UNIPROT:O55720; EMBL:AF003534; NID:g2738385; PIDN:AAB94431.1; PID:g2738385
C:Superfamily: Chilo iridescent virus hypothetical protein 032R
Alignment Scores:
Pred. No.: 0.122 Length: 602
Score: 122.00 Matches: 75
Percent Similarity: 38.94% Conservative: 57
Best Local Similarity: 22.12% Mismatches: 145
Query Match: 6.77% Indels: 62
DB: 2 Gaps: 10
US-10-089-688-1 (1-1053) x T03057 (1-602)
Qy 10 CCGTTTGGGAGTCTCAAAATCTCCAGCAGACATTGTGAAGATCTGAAG----- 60
Db 107 ProPheLeuProAsnPheMetLysProTyrAspLeuIleLysAsnValArgAlaAspPro 126
Qy 61 -----GAGAGCATGCTGTCTCGAAAGCAAGAC 90
Db 127 ArgGlnLysAsnProPheValThrLeuIleGluLysSerGlySerGluLysMetArg 146
Qy 91 ATTTCTGATAAAAGCAGAAAAGGCTACA-----GAGAAGTT 129
Db 147 LeuThrThrLysLysSerLysProLysThrPheSerAspMetAlaLeuPheGluTyrIle 166
Qy 130 TCCAAARAATCTGGTGGCCATGAAGAATCTGTATGGCCACAAATGAAGAAGACCTCAG 189
Db 167 AsnSerAsnMetThrLeuAlaGluLeuIleSerIleAsnValLysGlyLysArgGlnThr 186
Qy 190 ACAGAAGCAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGG-----CTCCTTAGC 243
Db 187 ArgValThrLeuAsnGluGluIleGlnLysIleLysAlaSerAspPheLysIleIleAsn 206
Qy 244 ACCCTGGTAGCTGATTACAGCTCATTTGAGGGCAAAAAGACGGTGGCTCAAAAT 303
Db 207 SerLeuLeuAsnGlnLeuMetIleAlaIleLeuIleAlaGlnLysLysLeuSerPheIle 226
Qy 304 -----TTCAACAATATTCTCAGAGACAAATTTGGTACGAGAACTCTACT 348
Db 227 HisAsnAspLeuHisPheAspAsnValLeuIleCysLysCysLeuGlnArgThr----Phe 245
Qy 349 GTTGAATACATCTGCCACCAACAGAAATATTGTTCATGTTATTGAAGGGTATGAATCT 408
Db 246 MetLeuTyrValPheGluTyrAsnSerIleCysTyrAlaLeuLeuProThrTyrGlyTyr 265
Qy 409 CCAGAATAGCTCTAAATTTGGGAATATGTTAAGAGAATGCATCAGACATGAACCACTT 468
Db 266 TyrProIleIleAspTyrGlyPheSerPheSerGluAspLeuIleGlyGlyProLeu 285
Qy 469 GCAAAATCATTTTGGTGGCAGACAGTTTATGATTTCTTCAGATATGTCGAATGTCA 528
Db 286 LeuThrGlyIleHisAsnAsnLysGlyTyrMetAsnHisGlnTyr----- 301

Qy 529 ACATTGTACATAGCTTCAGATGATTTGCCACATTCAGGATTTACTTACAGACATAAA 588
Db 302 -----AspGluPheThrAspPheLysThrMetLeuThrArgLeuSer 315
Qy 589 TTGCTCAGTCAGCAAAATTTTGGGAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAG 648
Db 316 TyrSerGlyTyrGlnPheGlyLeuAspLysLeuAspAlaPhe----- 329
Qy 649 TTACTTTCAATTCAGAAAATTATGTGCAAAAAGACAGTCACTG-----AAGCTTCTC 699
Db 330 -----GlnSerLeuIlePheAspLysLeuIle 338
Qy 700 GGTGAACACTACTAGATAGACACAACTTCCAAATTTATGACAAATATACATCACTTAACT 759
Db 339 SerLysLeuProIleAspLysGlnThrGlyTrpAspGluThrLysAspIleSerValSer 358
Qy 760 GAGAACCTCAAAATTAATGATGAACCTGCTCGAGAC-----AAAAGTCGCAACATC 810
Db 359 LysGlnLeuValArgHisIleArgIlePheValAspTyrLeuLysSerIleAsnArg 378
Qy 811 CAGTTTGGAGCCCTTTCACGTTTTTAAAGTGTGTTGTAGCCAACTCTAACAGACGACGCC 870
Db 379 Glu---SerPheGlnLysTyrAspTyrGluMetValAspMetIleGlySerLeuIle 397
Qy 871 ATCTAGACATCTCTCTCAAGAACCGCCAACTCATAGATTCCTCAGCAAGTTT 927
Db 398 IleLeuProLeuArgLysLysAsnThrGluAsnLeuValGluThrLeuAlaIlePhe 416
RESULT 10
D97848
hypothetical protein RC1188 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: D97848
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: D97848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-365 <KUR>
A:Cross-references: UNIPROT:Q92GD5; GB:AE006914; PIDN:AAL03726.1; PID:g15620317; GSPDB:G
C:Genetics:
A:Gene: RC1188
Alignment Scores:
Pred. No.: 0.286 Length: 365
Score: 117.00 Matches: 74
Percent Similarity: 36.31% Conservative: 60
Best Local Similarity: 20.05% Mismatches: 127
Query Match: 6.49% Indels: 108
DB: 2 Gaps: 14
US-10-089-688-1 (1-1053) x D97848 (1-365)
Qy 46 GTCAAGAACTCTGAAGGAGAGAGATGGTGTCTCGAAAAGCAAGACATTTCTGATAAAAAA 105
Db 60 IleSerAsnLeuLysGluThrAlaGluThrLysLeuLysHisTyrGluValGluGlnSer 79
Qy 106 GCAGAAAAGGCTACAGAGAAGTTTCCAAAAATCTGGTTGCCATGAAGAATTTCTGTAT 165
Db 80 IleGluLysAlaArgValGlnGluGluPheLysSerGlnArgAlaIleLysLeu 99
Qy 166 GCAGAAAATGAAGAGAGCTCAGACAGAGCAGTACTCAACTGCTCAAGAACTTAT 225
Db 100 AlaGluGluLysGlu-----IleAlaGlnArgIleAlaAlaLeuAsn 114
Qy 226 AAT-----AGTGGGCTCTTAGCACCCCTGGTAGCTGATTACAGCTCATTGACTTTGAG 279
Db 115 AsnLeuHisAsnGluPheIleLysAsnIleThrLysAspThrLysArgIleGlu---Glu 133

280 QY GGCMAAAAGACGTGGCTCAAATTTTCAACAATATTCTCAGAAGACAAATTTGGTAGCAGA 339
134 Db SerAsnLysArgLeuAspLysIleIleAsnLysGluLysGluAsnIle----- 150
340 QY ACTCCTACTGTTGAATACATCTGCACCAACAGAATATTTTGTTCATGTTATTGAAAGGG 399
150 Db ----- 150
400 QY TATGAATCTCCAGAAATAGCTCTAAATTGTGGAATAATGTTAAGAGAATGCATCAGACAT 459
151 Db -----IleAspHis 153
460 QY GAACCACTTCGCAAAATCATTTTGTGTCGCGAACAGTTTATGATTCTTTCAGATATGTC 519
154 Db GluGluLeuAsnArgGluIleLeuThrHisGluGluIlePheLysLeuAsnGln----- 171
520 QY GAATGTCAACATTTGCATAGCTTCAGATGCATTTGCCACATTTCAAGGATTTACTTACA 579
172 Db -----AlaTyrLysValLeuHisLys 178
580 QY AGACATAAATTTGCTCAGTCAGCAGAAATTTTGGAAACAGCATATTATGATAGATTTTTCAGTGAA 639
179 Db AsnHisLysLysIleThrGluGluHisLysGlnAlaHisThrGlu-----LeuAsnGlu 196
640 QY TATGAGAAGTTACTTCATCTCA-----GAAATATTATGTGCAAAAAGA 681
197 Db LeuAsnLysThrIleThrAsnLeuSerGlnGlnLeuLysLysGlyLeuThrProLys 216
682 QY CAGTCACCTGAAGCTTCTCGGTGAACACTACTAGATAGACACAACTTCACAAATTATGACA 741
217 Db LysValLysGluLeuLysGlyGluLeuGluPheHisGlnGluMetLeuLysIleHisLys 236
742 QY AAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAAC-----CTGCTGCGA 792
237 Db AlaHisValGluLysIleGluAsnSerLysLysIleValAspGlnGluIleLysTyr 256
793 QY GACAAAAGTCGCAACATCCAGTTGTGAGCGCTTTTCAGCTTTTAAAGTGTTGTGAGCCAAT 852
257 Db GluLysGluHisAsnIleAsnLysAspLysIleLysLysLeuGlySerAspIleLysGlu 276
853 QY CCTAACACAGACGACGCC-----ATCCTAGAC-----ATCCTCTCTCAAG 891
277 Db ArgTyrLysLysGluProGluLysTyrLeuAspAlaTyrGluLysTyrLeuValLeuLys 296
892 QY AACACGCGCAAACTCAT-----GAG 912
297 Db HisGlnHisLysAlaIleLysThrAspGlyValValLysAspIleGluHisHisAsnLys 316
913 QY TTCCTCAGCAAGTTTTCAGACGACGACGAGGATTTGATGAC----- 954
317 Db ValLeuAsnLysIleAsnValAspLysThrLysSerLeuAlaAsnLysIleAspGluGln 336
955 QY AGTTCCGTCACGACGACGAATTCCTCG-----GTCGATTTCACGC 993
337 Db ThrGlnValLysAspThrAsnThrArgSerLeuSerProAlaArgThrProAsnIleLys 356
994 QY GTT---AAACCGCGACGCGTGGGATC 1017
357 Db ValThrAsnAsnLysThrArgGlyIle 365

RESULT 11
A90551
Conserved hypothetical protein MYPU_3130 [imported] - Mycoplasma pulmonis (strain UAB CT
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 Hsequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: A90551
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: A90551
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-2819 <KUR>
A;Cross-references: UNIPROT_Q98QP8; GB:AL445566; PID:g14089727; PIDN:CAC13486.1; GSPDB:G-
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU 3130
A;Genetic code: SGC3

Alignment Scores:
Pred. No.: 0.466 Length: 2819
Score: 115.00 Matches: 83
Percent Similarity: 38.90% Conservative: 87
Best Local Similarity: 18.99% Mismatches: 127
Query Match: 6.38% Indels: 140
DB: 2 Gaps: 24

US-10-089-688-1 (1-1053) x A90551 (1-2819)

QY	13	TTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAGAAATCTG-----	57
Db	1404	PheGlyLysThr-----AenlleleSerAsnLeuTyAspLeuValLeu	1418
QY	58	-----AAGGAGAGCATGGCTGTCTGGAAAG--CAAGACATTCT	96
Db	1419	SerAlaLeuSerlleAsnGlnGluSerSerValLeuThrLysValGluAenlleThr	1438
QY	97	GATAAAAAGCAGAAAAGCGTACAGAA-----GAAGTTTCCAAAATCTGGTGCATG	150
Db	1439	HisLysAlaLeuAspGlnGlnAsnSerTyPheAlaValleLysThrLeuLeuSerVal	1458
QY	151	AAAGAAATCTGTATGGCAGCAAAATCAAAAGAGCGCTCAGACAGACAGCAGTGTCAACTT	210
Db	1459	LysGlu-----GlyAspValAspLysSerAspGluPheLysAlallelleLysLys	1475
QY	211	GCTCAAGAACTCTAT-----AATAGTGGCTCCTTAGCACCGCTGGTAGCTGATTTA---	261
Db	1476	PheLeuAsnLeuPheAlaLysAsnAspTyLeuLeuThrLysPheilleSerProIleVal	1495
QY	262	---CAGCTCANTGACTTTGAGGC-----AAAAAGACGTGGCTCAAAATTTCAACAAT	312
Db	1496	AspLysThrIleAspPheAspSerAlaThrLysSerSerilleValLysleulleLysAsn	1515
QY	313	ATTCTCAGACAGACAA-----ATTGGT	333
Db	1516	ValleuSerAspGluGlnAsnLeuLysPheSerAlaGlnlleleAspThrIlePheAla	1535
QY	334	ACGAGAACTCCTACTGTGTAATAC-----ATC	360
Db	1536	AsnLysAsnLysTyLeuAlaTyLeuAsnLeuAsnAspPheleulleAspAlaLeuVal	1555
QY	361	TGCACCCACAGAAATTTGTTTCATG-----	387
Db	1556	AspAsnLysGluAsnilleleAsnLeuSerSerAsnThrLeuGlyLyslleLysThrAsp	1575
QY	388	-----TTATTGAAAGGATGATGAATCTCCAGAAATAGCTCTAAATTTGGGA	432
Db	1576	AspGluPheLysAsnilleleLysSerPheIleThrLysAsnLeuMetThrSerAspSer	1595
QY	433	ATAATGTTAAGAAATGATCATCAGA-----CATGAACCACTTGCMAAA	474
Db	1596	SerIleThrGluGluGluIleAsnThrIleValLysSerAlaHisGluLeuIleSer---	1614
QY	475	ATCATTTTGTGTGCGACAGTTTAT-----GATTTCTTCAGATATGTC	519
Db	1615	---IleValAspSerSerAsnPheLeuAenGlnLeuLysAspPhePheAlaLeuLeu	1633
QY	520	GAAATGTCAACATTTGCATAGCTTCAGATGCA-----TTTGCCACATTTCAAGGATTTA	573
Db	1634	GluLysAsnLysLysAspSerLysSerThrSerLeuGluPheSerThrPheMet-----	1651
QY	574	CTTACAAGACATAAATGTCTAGTCAGCAAAATTTTGGAAACAGCATATGATAGATTTTC	633
Db	1652	-----LvslsSerilleSerGluPheLeuAspAlaGlu---AsnAsnTyLeu	1666

Qy 634 AGTGAATATGAGAAGTTACTTCATTCTAGAAAAATTATGTGACAAAGACAGCAGTGCACTGAAG 693
:: : :: :: :: :: :: :: ::
Db 1667 AlaleuPheLysValLeuSerAsnLeuLeuSerThrGlnSer----- 1682

Qy 694 CTTCTCGGTGAAGTAATACTTAGATAGACACAATTCCACA-----ATTATG 738
||||| :: :: |||||
Db 1683 -----AppGlnSerPheThrLysSerIleSerSerIleLeu 1695

Qy 739 ACAAATATCATCATGAACCTCAAACTTATGATGAACCTGCTGCAGACAAA 798
::::: ||| ::
Db 1696 ArgValPheLeuArgLysGluLysAlaLeuGluLeuMet-----LeuLysLysGln 1712

Qy 799 AGTCGGCAACATCCAGTTT-----GAGGCCCTTTCACGTTTTTAAAGGTG 840
||||| :::
Db 1713 PheArgSnPheGluLeuHisAaspIleLysValAspAlaIleLeuLys--- 1731

Qy 841 TTTGTACCCTCTAACACAGCGCACCCCTCTAGACATCTCTCCACAGAAC----- 894
||||| :: ::
Db 1732 PheIlePheArgAsnGlnValIleAaspPheIleGluThrIleLeuAsnArgValVal 1751

Qy 895 -----CAGGCCAAACTCATAGAGTTCCCTC 918
||||| :: ::
Db 1752 SerGlnLysGluGluTySerLysLeuAsnSerTyrglGluLeuLeuTyLysPheLeu 1771

Qy 919 AGCAAGTTTCAGACAGCAGCGAGATGTATGAGCAGTTCCGTA----- 963
||||| :: ::
Db 1772 SetSerAsnLysAsnGluThrLeuAsnPhePhelyLysGlnIleGlyGluLeuLysAsn 1791

Qy 964 -----CCGACGACGAATCCCGGTCGATTACGCGTTAAA 999
||||| :: ::
Db 1792 SerLysLeuLeuAsnProLeuAlaLysSerPheIleGluSerGluLeuLys 1808

RESULT 12

T18372 repeat organellar protein - Plasmidium chabaudi
C;Species: Plasmidium chabaudi
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18372
R;Werner, E.B.; Taylor, W.R.; Holder, A.A.
Mol. Biochem. Parasitol. 94, 185-196, 1998
A;Title: A Plasmidium chabaudi protein contains a repetitive region with a predicted spe
A;Reference number: Z18922; MUID:98418765; PMID:9747969
A;Accession: T18372
A;Status: preliminary; translated from GE/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1939 <WER>
A;Cross-references: UNIPROT.Q25662; EMBL.U43145; NID:g1151157; PID:g1151158; PIDN:AAC634

Alignment Scores:

Pred. No.: 0.593 Length: 1939
Score: 113.50 Matches: 57
Percent Similarity: 38.78% Conservative: 57
Best Local Similarity: 19.39% Mismatches: 113
Query Match: 6.30% Indels: 67
DB: 2 Gaps: 11

US-10-089-688-1 (1-1053) x T18372 (1-1939)

Qy 52 AATCTGAAGAGAGCATGGCTGTTCTGGAAAACAGACACATTTCTGATAAAAAGCAGAA 111
||||| :: :: ::
Db 1433 AsnLeuAsnLysAsnValGluAspLysThrGlyAspAspileAsnCysGluLysAsnAsn 1452

Qy 112 AAGCCTACACAAAGAGTTCCC-----AAAAATCTGGTGGCCATGAAGAAATCTGTAT 165
:: :: :: :: ::
Db 1453 AspGlnAlaLysGluIleSeryLeuLysAspGluIleLysLysIleSerMetLeuTy 1472

Qy 166 GGC-----ACAAATGA AAAAGCGCTTCACACAGACAGCAGTAGCTCAACTTGCTCAAGAA 219
||||| |||||
Db 1473 GlyGluCluleuAsnArgLysAsnSerTyrglGluLysValLysAsnLeuThrAsnGlu 1492

Qy 220 CTCTATAATAGTGGGCTCTCTAGACCGCTGGTACGTGATTTACAGCTCATTTGAG 279
|||
Db 1493 Leu-----LvsgLuLeuLysIleAcqAsnLysLvs 1502

Qy	280	GGCAAAAAGACGTGGCTCAAAATT-----	303
		: : : : : : :	
Db	1503	GlyGluGluAlaIleAlaGluLeuAsnLysLeuLysA	1522
		: : : : : : :	
Qy	304	-----TTCAACAATAATTCACAGACACAAAT	339
		: : : : : : :	
Db	1523	ValLysGlnAsnAspGluSerSerAsnAsnLysLeu	1541
		: : : : : : :	
Qy	340	ACTCTACTGTGAATACATCTGCACCAACAGAAAT	384
		: : : : : : :	
Db	1542	ThrPro-----GluTyrValSerAsnAspAspLys	1559
		: : : : : : :	
Qy	385	ATGTTATGTAAGGGTATGAAATCTCCAGAAATAG	444
		: : : : : : :	
Db	1560	LeuValLeuLysLeuLysGluLysProAspLeuT	1579
		: : : : : : :	
Qy	445	GNATGCATCAGACATGAACACCTTGCAAA-----	498
		: : : : : : :	
Db	1580	GluAsnPheArgValMetSerIleValLysGluAsn	1599
		: : : : : : :	
Qy	499	TATGATTCTTCAGATATGTCGAAATGTCACATTT	558
		: : : : : : :	
Db	1600	ValGlyIleTyrSerTyrPheLysLysCysGluLys	1619
		: : : : : : :	
Qy	559	ACATTCAAGGATTACTTACACAGACATAAATTC	618
		: : : : : : :	
Db	1620	Ile-----CysLeuValLeuLysAspIleLeuSer	1636
		: : : : : : :	
Qy	619	TATGATAGATTTTTCAGTGAATATGAGAGTTACT	678
		: : : : : : :	
Db	1637	PheValAsnLeuPheGluLysIleAspLysIleLeu	1656
		: : : : : : :	
Qy	679	-----ACACAGTCACTG	690
		: : : : : : :	
Db	1657	GluIleArgIleLeuPheLeuArgTyrPheSerPhe	1676
		: : : : : : :	
Qy	691	AAGCTTCTCGGTGAATCTACTACTAGAT-----	738
		: : : : : : :	
Db	1677	LysCysValAsnGluGluTyrValAsnAsnGluArg	1696
		: : : : : : :	
Qy	739	ACAAATATACATCAGTAACTCGAGACCTCAAAAT	780
		: : : : : : :	
Db	1697	GlnThrTyrLeuGluThrAlaSerAsnLeuLysGlu	1710
		: : : : : : :	

RESULT 13

G70163

hypothetical protein BB0512 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: G70163

R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hat Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: G70163

A>Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-2166 <KLE>

A:Cross-references: UNIPROT:O51465; GB:AE0001153; GB:AE000783; NID:g2688419; PID: A:Experimental source: strain B31

Alignment Scores:		
Pred. No.:	0.653	Length: 2166
Score:	113.00	Matches: 80
Percent Similarity:	43.30%	Conservative: 72
Best Local Similarity:	22.79%	Mismatches: 116
Query Match:	6.27%	Indels: 84
DB:	2	Gaps: 21

Alignment Scores:		
Pred. No.:	0.653	Length: 2166
Score:	113.00	Matches: 80
Percent Similarity:	43.30%	Conservative: 72
Best Local Similarity:	22.79%	Mismatches: 116
Query Match:	6.27%	Indels: 84
DB:	2	Gaps: 21
	--	

A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.
A:Reference number: A70100; PMID:98065943; PMID:9403685
A:Accession: G70163
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2166 <XLE>
A:Cross-references: UNIPROT:O51465; GB:AF001153; NID:G2688419; PID:G2688419
A:Experimental source: strain B31

C/SpecIES: Borrelia burgdorferi (Lyme disease spirochete)
 C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
 C/Accession: G70163
 R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra,
 son, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hat
 Nature 390, 580-586, 1997

Qy	691	AAGCTTCTCGGTGAAGCTACTACTAGATG	-----AGACACAACTTTCACAAATTATG	738
Db	1677	LysCysValaAsnGluGluTyrValaAsnGlu	ArgTyrGluTyrSerTrpAlaLeuPhe	1696
Qy	739	ACAAATACATCAGTAAACCTGAGACCTCA	ATAATTATGATG	780
Db	1697	GluThrTyrLeuGluThrAlaSerAsnLeuLys	GluMet	1710

RESULT 13

C780C1

Qy		619	TATGATAGATTTTTCAGTGGAATATGAAAGCTTA	CCTCAATTGGACAAA	678
	:	:	:::	:::	:
Db		1637	PheValAsnLeuPheGluLysIleAspLysIleLeuTrpLysGlnMetYrIleProThr		1656
	:	:	:::	:::	:
	:	:	:::	:::	:
Qy		679	-----AGACAGCTCACTG	690	
	:	:	:::	:	:
Db		1657	GluIleArgileLeuPheLeuAryrPheSerPheLeuAspLysLeuAgsenTyVal		1676

D5	1360	GLUASNPHENALYVALMETSERILEVALLYSGLUKUNILYASASNVALGILUASAPPLYLSILE	1357
Qy	499	TATGATTCTTCAGATATGTCGAATTTGCAACATTTTGACATGCTTCAGATGCATTTGCC	558
Db	1600	VALGLYLILETYRSERTYRPHYLYSYSCYSGLULYSGLULEULYASASNAPMETLEUVAL	1619
Qy	559	ACATTCAAGATTTACTTCAAGACATAAATTTCTCAGTCAGCAAAATTTTTTGGACAGCAT	618
Db	1620	Ile-----CysLeuValLeuLysASPILLESerIleLeuPheLeuAsnASPASP	1636

1542	ThrPro-----GluTyrValSerAsnAspAspLeuGlnIysAspTrpLysAlaAsn	1559
385	ATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGGAAATAATGTTAAGA	444
1560	LeuValLeuLysLeuLysGlnIysProAspLeuTrpAspAsnIleAsnSerLeuGlnLys	1579
445	GAATGCATCAGACATGAACCACTTGCAGAAA-----ATCATTTTGTGGTCGGAACACGTTT	498
1590	GlnAlaPheLeuIleValMetGluLeuValGluAlaAsnIleAsnValGlnIleLeuValGlnAla	1600

[illegible]

-1.rpr

QY 280 GGCAAAAAGACCTGGCTCAATT----- 303

Page 9

Db 398 PheLysIleLysSerGluSerPheLysLeuSerPheLysLeuLysGluLysGlyPhe 417
QY 556 GCCACATTCAGGATTACCTTACAGACATATAAATGCTCAGTGCAGAAATTTTCGAACAG 615
Db 418 LeuGluLeuGluLeuLeuLeuLeuGlnGlyLysGluLeuLeuLeuLeuLeuLeuLeuLeu 437
QY 616 CATTATGATAGATATTTTC----- 633
Db 438 GluLeuGluLeuPheSerLeuLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 457
QY 634 -----ACTGAATATGGAAGTACTT 654
Db 458 LysLysIleGluLysProSerProLysProLysGluSerAlaPheIleProLeuLeu 477
QY 655 CATTGAGAAATATGTCACAAAAGACAGTCACTGAGCTTCCGGTGAAGTACTACTA 714
Db 478 GluSerGluGly-----IleGlyArgProSerThrTyrAlaSerPheLeuAspLeuLeu 496
QY 715 GATAGACACAACTTCACAAATTATGACAAA----- 744
Db 497 LysArgLysTyrIleSerIleAspThrLysThrAsnAlaIleThrProThrSerGlnGly 516
QY 745 -----TACATCAGTAAACCTCAGAACCTCAAAATTAATGATGAACTGCTG 789
Db 517 LeuGluValIleSerPhePheLysAspLysGluValAspPheIleAlaLeuThrSer 536
QY 790 CGAGACAAAAGTCCGACATCCAGTTTGAGCCCTTTCACGTTTTTAAGGTGTTGTAGCC 849
Db 537 LysAspLysSerLys-----LeuGly 543
QY 850 AATCCTAACAGACG---CAGCCCATCTAGACATCTCTCAGAACACCGCCAAA--- 903
Db 544 AsnThrThrLysGlnPheGluGluCysLeuAspLeuIleMetArgGlyGluAlaSerTyr 563
QY 904 -----CTCAGAGTTCTCAGCAAGTTTCAGAACGAC-----AGGACG 942
Db 564 GluLysPheMetLeuGluValIleSerLysLeuLysSerThrAlaLysPheTyrGlnThr 583
QY 943 GATTGTATGACGATTCGTCACGACGACGAAATTCGCGGTGATTTACGCTTAACCG 1002
Db 584 GlnSerThrAspAsnMetProThrAspLys-----GlnLeuGluLeuIleAsp 600
QY 1003 CGGACGGTGGATCGGATTCAGACGACCGACCTCAGCNA 1044
Db 601 LysIleCysLysAspLysLysLeuGlnLysProSerGlnGlu 614

RESULT 15

T18489 hypothetical protein C0820w - malaria parasite (Plasmodium falciparum)

C/Species: Plasmodium falciparum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T18489

R/Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A/Reference number: Z18935

A/Accession: T18489

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4981 <LAW>

A/Cross-references: UNIPROT:O77372; EMBL:Z98551; NID:e1331903; PID:e1331910; PIDN:CAB111

C/Genetics:

A/Map position: 3

A/Note: C0820w

Alignment Scores:

Pred. No.:	0.984	Length:	4981
Score:	111.00	Matches:	59
Percent Similarity:	36.73%	Conservative:	49
Best Local Similarity:	20.07%	Mismatches:	104
Query Match:	6.16%	Indels:	82
DB:	2	Gaps:	10

US-10-089-688-1 (1-1053) x T18489 (1-4981)

QY 298 CAAATTTTCAACATATTTCTCAGACGACAAATTTGGTACGAGAACTCTCTACTGTTGAATAC 357
Db 1521 LysIlePheLeuAsnTyrIleArgLysGluLysGluHisIleTyrAsnPheGlnThrTyr 1540
QY 358 ATCTGCACCCCAACAGAAATATTTTTCATGCTTTTGAAGGGTATGAATCT- 408
Db 1541 IleThrThrPhePheThrLeuLeuGluMetLeuLysAsnIleGlnSerAspLysAsp 1560
QY 409 ---CCAGAAATAGCTCTAAATTTGGAATAATGTTAAGAAATGTCATCAGACATGAACCA 465
Db 1561 TyrGluLysThrLeuTyrAsnIleGlyIleIleAsnMetMetCysValAlaValLysPhe 1580
QY 466 CTTGCA-----AAATCATTT- 480
Db 1581 LeuLysGluLysCysProAsnLysIleIleSerIlePhePheLysLeuLeuGluVal 1600
QY 481 -----TTGTGGTCCGAAACAG-----TTTATGATTTCTTCAGATAT 516
Db 1601 ProIleProLeuGluGluGluAspLysValValThrLeuTyrLysAsnTyrSerTyr 1620
QY 517 GTCGAAATGTCACATTTGATAGTCTTCAGATGTCATTTGCCACATTCAGAGATTTACTT 576
Db 1621 IleTyrArgThrHisPhe-----PheAsnLeuTyrLysAspLysCys 1634
QY 577 ACAAGACATAAATTTGCTCAGTCAGAAATTTTGGAAACAGCATTATGATAGATTTTTCAGT 636
Db 1635 Thr-----ThrLeuTyrThrLysAsnCysMet 1643
QY 637 GAATATGAGAGTTACTTTCATTCAGAAAATATGTGACAAAAGACAGCTCACTGAAGCTT 696
Db 1644 GluTyrIleLysMetMetArgAsnAsnIleTyrLeuAsnLysLysLysAsn----- 1660
QY 697 CTCGGTGAATCTACTAGTAGACACAACTTCACAAATATGACAAATACATCAGT--- 753
Db 1661 ---AspSerGlnValAspPheAsnAsnPheIleIleLeuSerAsnMetTyrAsnAsn 1678
QY 754 -----AACTGAGAACCTCAAAATTAATGATGAACCTGCTGCGAGACAAAGT 801
Db 1679 SerAspThrPheLysLysLysAsnIleGlnAsnValLeuGlnTyrLeuAsnAspTyrGln 1698
QY 802 CGCAACATCCAGTTTGAGGCTTTTCAGTTTAAAGTGTGTTGAGCAAT----- 852
Db 1699 GlnValTyrLysLysAsnValValHisIlePheAsnMetTyrIleSerAsnValThr 1718
QY 853 -----CCTAACACAGACG 864
Db 1719 HisLysGluMetGluAsnGlyMetTyrArgLysGluGlyAsnLysArgGluAsnAsnThr 1738
QY 865 CAGCCCATCTAGACATCTCTCTCAAGAACAG----- 897
Db 1739 TyrAspLysIleAsnIleAspAspLysAsnAsnPheGlyGlnProIleTyrLysTyrVal 1758
QY 898 GCCAAACTCATAGAGTTCTCCACCAAGTTTCAGAACGACGAGGAGGATGTTATGAGCAGT 957
Db 1759 AspLysLeuThrSerLysArgAsnThrLeuAspAsnSerSerIleAsnCysLeuGluSer 1778
QY 958 TCCGTACCGACGACGAATTCCTCGGGTCGATTTCACGGTTAAA 999
Db 1779 SerTyrValThrLysGlySerAsnThrLysIleArgIleLys 1792

Search completed: November 10, 2004, 20:52:59

Job time : 74.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 10, 2004, 20:33:32 ; Search time 236.5 Seconds
(without alignments)
5123.632 Million cell updates/sec

Title: US-10-089-688-1

Perfect score: 1803

Sequence: 1 atgcgcgtccgtttgggaa.....cagctcagcaagagcttaa 1053

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPRO.spool.p/US10089688/runat.10112004.152246.4843/app.query.fasta.1.1223
-DB=Uniprot.02 -OPMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOPEXT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US10089688 @CGN 1.1.244 @runat.10112004.152246.4843 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot.02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1635.5	90.7	341 1	CB39 HUMAN Q9Y376 homo sapien
2	1619.5	89.8	341 2	Q8VDZ8 mus musculus
3	1614.5	89.5	341 1	CB39 MOUSE Q06138 mus musculus
4	1614.5	89.5	341 2	Q6IP72 xenopus lae
5	1614.5	89.5	341 2	AAH72045 xenopus l
6	1590.5	88.2	341 2	Q724X0 Q6IQ12 brachydanio
7	1537.5	85.3	341 2	AAH71393 Q6PAB4 xenopus lae
8	1537.5	85.3	341 2	AAH71393 Q6PAB4 xenopus lae
9	1339.5	74.3	337 2	AAH60384 Q9DB16 mus musculus
10	1339.5	74.3	337 2	AAH60384 Q9DB16 mus musculus
11	1298.5	72.0	337 1	C39L MOUSE Q9H944 homo sapien
12	1297.5	72.0	337 1	C39L HUMAN Q9H944 homo sapien
13	1287.5	71.4	337 2	AAQ93064 Q803V8 brachydanio
14	1287	71.4	343 2	Q803V8 Q7PF07 anopheles g
15	1163.5	64.5	332 2	Q7PF07 anopheles g
16	1163.5	64.5	335 2	Q7PF07 anopheles g

17	1123	62.3	339	1	MO25 DROME
18	1123	62.3	339	2	AAW5031
19	1054.5	58.5	377	2	Q7Z2A5
20	1054.5	58.5	636	2	Q21643
21	1006	55.8	338	1	MO2M CABEL
22	944.5	52.4	205	2	Q8K312
23	766	42.5	329	1	YFV6 SCHPO
24	764.5	42.4	321	1	DE76 CHLPR
25	726.5	40.3	390	2	Q6CAN7
26	711.5	39.5	343	1	MO2N ARATH
27	711.5	39.5	343	2	AA116128
28	705.5	39.1	343	1	MO2M ARATH
29	705.5	39.1	343	2	AA106959
30	686	38.0	338	2	Q6BSQ5
31	669.5	37.1	345	2	Q8L9L9
32	651.5	36.1	337	2	Q7X1Q4
33	623	34.6	384	1	HYMA EMENI
34	614	34.1	348	1	MO2L ARATH
35	614	34.1	348	2	AA24657
36	584	32.4	370	2	Q873K5
37	571	31.7	364	2	Q6CWP5
38	560	31.1	362	2	Q757C0
39	560	31.1	362	2	AA52777
40	511	28.3	422	2	Q6FWG7
41	501	27.8	399	1	HYM1 YEAST
42	455.5	25.3	332	2	Q7QTG6
43	327	18.1	103	2	Q8K038
44	283.5	15.7	509	2	Q7YYL6
45	140.5	7.8	1111	2	Q9U0K5

ALIGNMENTS

RESULT 1

CB39_HUMAN ID CB39 HUMAN STANDARD; PRT; 341 AA.

AC Q9Y376; DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Calcium binding protein 39 (Mo25 protein) (CGI-66).
GN Name=CB39; Synonyms=Mo25;
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;

RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=10810093;
RA Jin W., Shi J., Ren S., Gu J., Fu S., Huang Q., Dong H., Yu Y., Fu G.,
Wang Y., Chen Z., Han Z.;

RT "A novel gene expressed in the human hypothalamus.";
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RX MEDLINE=2238257; PubMed=10810093;
RA Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Result No. Score Query Length DB ID Description

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4].
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 10-339.
 RX PubMed=14730349; DOI=10.1038/nmsb716;
 RA Milburn C.C., Boudreau J., Deak M., Alessi D.R., van Aalten D.M.;
 RA "Crystal structure of Mo25 alpha in complex with the C terminus of the
 RT pseudo kinase STE20-related adaptor.";
 RL Nat. Struct. Mol. Biol. 11:193-200(2004).
 CC -!- FUNCTION: Together with the STE20-related adaptor-alpha (STRAD
 CC alpha) pseudo-kinase, forms a regulatory complex capable of
 CC stimulating the activity of SKL1.
 CC -!- SIMILARITY: Belongs to the Mo25 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF151824; AAD34061.1; -;
 DR EMBL; AF113536; AAF14873.1; -;
 DR EMBL; BC020570; AAH20570.1; -;
 DR PDB; 1UPK; X-ray; A=1-341.
 DR InFAct; Q9Y376; -;
 DR Genew; HGNC:20292; CAB39.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 KW 3D-structure.
 SQ SEQUENCE 341 AA; 39869 MW; EC710A528B6F9811 CRC64;

Alignment Scores:

Pred. No.: 6,78e-114 Length: 341
 Score: 1635.50 Matches: 329
 Percent Similarity: 94.29% Conservative: 1
 Best Local Similarity: 94.00% Mismatches: 11
 Query Match: 90.71% Indels: 9
 DB: 1 Gaps: 2

US-10-089-688-1 (1-1053) x CB39_HUMAN (1-341)

QY 1 ATGCGTTCCTCCGTTGGGAGTCTCAAAATCCAGCAGACATTTGTGAAGAATCTGAAG 60
 DB 1 MetProPheProPheGlySerHisLysSerProAlaAspIleValLysAsnLeuLys 20
 QY 61 GAGAGCATGGCTGTTCTTGAAAGCAAGACATTTCTGATAAAGAGCGAGAAAGCTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAAGTTCCAAAATCTGTCCTGCAAGAAATCTCTATGGCAAAATCAAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysAsnSerGlyLeuLeu 80
 QY 241 AGACCCCTGGTACTGATTTACAGCTCATTTGACTTTGAGGCGCAAAAAGACCTGGCTCAA 300
 DB 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100

RESULT 2

Q8VDZ8 Q8VDZ8 PRELIMINARY; PRT; 341 AA.
 AC Q8VDZ8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cab39 protein.
 GN Name=Cab39;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

QY 301 ATTTTCAACAATATTCTCAGAGACAAATTTGTCAGAGACCTCTACTGTTGAATACATC 360
 DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
 QY 361 TGCACCCCAACAATATTTTGTTCATGTATTGAAGGGTATGAATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 140
 QY 421 CTAAATTTGGGAATATGTTTAAGAGAAATGCATCAGACATGAACCACTTGCAGAAATCAT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
 QY 481 TTGTGTCGGAACAGTTTATCATTTCTTCAGATATGTGCAAAATGTCAACATTTGCACATA 540
 DB 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGCAATTTGCCACATTCAGAGATTACTACAGACATATAAATTTGCTCAGTCGA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 QY 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
 DB 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
 QY 661 GAAATATTGTGACAAAAGACAGTCATCAGAGCTTCTCGTGAACTACTACTAGATAGA 720
 DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyLeuLeuLeuAspArg 240
 QY 721 CACAACCTTCAATATTGACAAATACATCAGTAACCTGAGAACCTCAAAATTAATGATG 780
 DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 QY 781 AACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTTCAGGCGCTTTCACGTTTTTAAGGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCCAATCTTAACAGACGCGAGCCCATCTAGACATCTCTCTCAGAACCCAGGCC 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
 QY 901 AAACCTCATAGATTCTCTCAGCAAGTTTCAGAACGACGAGGACGATTTGTATGAGCAGTTCC 960
 DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTACCGACGACGAATTCCTCCGGTTCGATTTACGCGTTTAAACCGCGGACGCTGGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 QY 1021 GATTTGAAGACGACGAGCTCAGCAAGAGCT 1050
 DB 332 AspLeuLysArgProAlaGlnGlnGluAla 341

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 Kzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 Jones S.J., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 (2)
 SEQUENCE FROM N.A.
 STRAIN-CZECH II;
 TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 R. Strauberg R.;
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 EMBL; BC020041; AA020041.1; -;
 MGD; MGI:107438; Cab39
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39843 MW; E7FECA529D6FE811 CRC64;

Alignment Scores:

Pred. No.: 1.07e-112 Length: 341
 Score: 1619.50 Matches: 326
 Percent Similarity: 93.43% Conservative: 1
 Best Local Similarity: 93.14% Mismatches: 14
 Query Match: 89.82% Indels: 9
 DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x Q8VDZ8 (1-341)

QY 1 ATGCGCTCCGTTGGGAAGTCTCACAAATCTCCAGCAGACATTTGTGAAGATCTGAAG 60
 DB 1 MetProPheProPheGlyLysSerHisLysSerProHisAspLeuValLysAsnLeuLys 20
 QY 61 GAGACATGCTGTTCTGGAAGCAGACATTTCTGATATAAAGCAGAAAGCGGTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspLysSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAGTTTCCAAAATCTGGTTCATCAAGAAATCTGTATGTCACAAATGAAGA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLeuTyrGlyThrAsnGluLys 60
 QY 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
 QY 241 AGCACCTCGTAGCTGATTACAGCTCATTTGACTTTGAGGGCAAAAGCGGTCTCA 300
 DB 81 GlyThrLeuValAlaAspLeuGlnLeuLeuAspPheGluGlyLysLysAspValAlaGln 100
 QY 301 ATTTTCACATATTTCTCAGACAGCAATTTGGTACGAGAACTCTACTGTTGATATCATC 360
 DB 101 IlePheAsnAsnLeuArgGlnLeuGlyThrArgThrProThrValGluTyrIle 120
 QY 361 TGCACCAACAGAAATATTTGTTTCATGTTAAGAGGGTATCAATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProGluIleAla 140
 QY 421 CTAAATTTGGAATATTTGTTAAGAGATGTCATCAGACATCAACCTTGTGCAAAATCAT 480
 DB 141 LeuAsnGlyGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
 QY 481 TTGTGGTCCGAACAGTTTTATGATTTCTTCAGATATGTCGAATGTCACATTTGCATA 540

DB 161 LeuTrpSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATCATTTGTCACATTCAGAGATTACTTACAGACATAAATTTGCTCAGTGCA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 QY 601 GAATTTTGAACAGCAGATTATGATAGATTTTTCAGTGAATATGAGAAAGTTACTTCA 660
 DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 QY 661 GAAATTTATGTCACAAAGACAGTCACTGAAGCTTCTCGTGAACCTACTACTAGATAGA 720
 DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAspArg 240
 QY 721 CACAACCTTCAATATATGACAAATATCATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
 DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 QY 781 AACCTGCTGGAGACAAAGTCGCAACATCAGTTTGGAGCCCTTCACTGTTTAAAGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCCAATCTTAACAGACGACGAGCCCATCTAGACATCTCTCAAGAACCGCC 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLysAsnGlnThr 300
 QY 901 AAACCTCATAGATTCTCTCAGCAAGTTTTCAGAACGACGAGCAGGATTTGATGAGCAGTCC 960
 DB 301 LysLeuLeuGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTACCGACGACGAATTTCCGGGTGATTTACGGCTTAACCGCGGAGCGCTGGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 QY 1021 GATTTCAAGACAGCAGCTCAGCAAGAGCT 1050
 DB 332 AspLeuLysArgAlaAlaGlnGlnGluAla 341

RESULT 3
 CB39 MOUSE
 ID CB39 MOUSE STANDARD; PRT; 341 AA.
 AC Q06138;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Calcium binding protein 39 (Mo25 protein).
 GN Name-Cab39; Synonyms=Mo25;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93119656; PubMed=8418809;
 RA Miyamoto H., Matsushiro A., Nozaki M.;
 RT "Molecular cloning of a novel mRNA sequence expressed in cleavage
 RT stage mouse embryos."
 RL Mol. Reprod. Dev. 34:1-7(1993).
 CC -I- FUNCTION: Together with the STR20-related adaptor-alpha (STRAD
 CC alpha) pseudokinase, forms a regulatory complex capable of
 CC stimulating the activity of SRK11.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -I- DEVELOPMENTAL STAGE: Transcribed during early mouse development.
 CC Detected at all developmental stages from the egg through the
 CC blastocyst, most abundant at the 2-cell stage.
 CC -I- SIMILARITY: Belongs to the Mo25 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; S51858; AAB24801.1; -;
DR MGD; MGI:107438; Cab39.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39842 MW; E7F668529D6FE811 CRC64;

Alignment Scores:

Pred. No.: 2,52e-112 Length: 341
Score: 1614.50 Matches: 325
Percent Similarity: 93.43% Conservative: 2
Best Local Similarity: 92.86% Mismatches: 14
Query Match: 89.55% Indels: 9
DB: 1 Gaps: 2

US-10-089-688-1 (1-1053) x CB39_MOUSE (1-341)

```
QY 1 ATGCCGTTCCCGTTGGAGTCTCAAAATCTCCAGCAGACATTGTGAAGATCTGAAG 60
DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGCTGTTCTCGGAAAGCAAGACATTCTGATAAAAGCAGAAAAGGCTACA 120
DB 21 GluSerMetAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAAGAAGTTTCCAAAATCTGGTTCATGATGCAAGAAATCTGTATGCGCACAAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluIleLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCTCAGACAGACAGTAGCTCAACTGCTCAAGACTCTATATAGTGGCTCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGCACCTGGTGTGATGTTACACTCATGCTTTCAGGCGCAAAAGACGTCGCTCAA 300
DB 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 100
QY 301 ATTTTCAACATATTTCTCAGAAAGCAAAATGGTACGAGAACTCTCTGTTGAATACATC 360
DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCACCCACAGATATTTGTTTCATGTTATTAAGAGGTATGATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGTGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAAAATCAT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle 160
QY 481 TTGTGGTCGGAACAGTTTATGATTCTTCAGATATGTCGAAATGTCAACATTTGACATA 540
DB 161 LeuTyrSerGluGlnPheTyrAspPheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATGATTGTCACATTCAGGATTTACTTACAGACATAAATGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
QY 601 GAAATTTTGGAAACAGCATTATGATAGATTTTTCAGTGAATATGAGAACTTACTTCA 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATTTATGTACAAAAGACAGTCACTGAGCTTCTCGGTGAACACTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAACCTTCAATATACAAAATACATCATGTAACCTGAGAACTTCAAAATTAATCATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACCTGCTCGCAGACAAAAGTCGCAACATCCAGTTTGAGGCTTTTACGTTTTTAAGGTG 840
```

```
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCATCTTAACAGAGCGAGCCCATCTCTAGACATCTCTCAAGAACAGGCGC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnThr 300
QY 901 AAACCTCATAGAGTTCTCTCAGCAAGTTTTCAGAACGACGACGACGAGTGTATGAGCAGTTCC 960
DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
QY 961 GTACCGACGACGAATTCGCGGTTCGATTTCGCTTAAACCGCGGACGCGGTGGATCAGG 1020
DB 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
QY 1021 GATTTGAAAGACGACGCTCAGCAAGAGCT 1050
DB 332 AsnLeuLysArgAlaAlaGlnGlnGluAla 341

RESULT 4
Q6IP72 PRELIMINARY; PRT; 341 AA.
AC Q6IP72;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE MGC78903 protein.
GN Name=MGC78903;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S.L., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072045; AAH72045.1; -.
DR InterPro; IPR008938; ARM.
```



```
DR InterPro: IPR004892; Mo25.
DR Pfam: PF03204; Mo25; 1.
SQ SEQUENCE 341 AA; 39777 MW; 5AFEAF9325BB2F16 CRC64;

Alignment Scores:
Pred. No.: 2,52e-112 Length: 341
Score: 1614.50 Matches: 323
Percent Similarity: 93.71% Conservativeness: 5
Best Local Similarity: 92.29% Mismatches: 13
Query Match: 89.55% Indels: 9
DB: 2 Gaps: 2

US-10-089-688-1 (1-1053) x Q6IP72 (1-341)

QY 1 ATGCGTCCGTTGGAGAGCTCTCAAAATCTCCAGCAGACATTTGGAAGATCTGAAG 60
DB 1 MetProPheProPheGlyLysSerHisLysSerProAlaAspIleValLysAsnLeuLys 20
QY 61 GAGAGCATGGCTGTTCTGGAAGACAGACATTTCTGATRAAAAGCAGAAAGGCTACA 120
DB 21 GluSerIleAlaValLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaThr 40
QY 121 GAGAAAGTTTCCAAAATCTGGTTGCCATGAAAGAAATTTCTGTATGGCACAATAATGAAAA 180
DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLysLeuTyrGlyThrAsnGluLys 60
QY 181 GAGCTCAGACAGACAGTAGTCTCAACTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyrAsnSerGlyLeuLeu 80
QY 241 AGACCTCGTAGTGTATTTACAGCTCATTGACTTTGAGGGCAGAAAGAGCGTGCCTCA 300
DB 81 GlyThrLeuValAlaAspLeuGlnLeuIleAspPheGluLysLysAspValAlaGln 100
QY 301 ATTTTCAACAATATTTCTCAGAGACAAATTTGGTACGAGAACTCTCTACTGTGTGAATACATC 360
DB 101 IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrIle 120
QY 361 TGCAACCAACAGAAATATTTTGTTCATGTTATTTGAAGGGTATGAATCTCCAGAAATAGCT 420
DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLysGlyTyrGluSerProGluIleAla 140
QY 421 CTAAATTTGGTAATATTTAAGAAATGATCATCAGACATCAACCACTTGCAGAAATCAATT 480
DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
QY 481 TTGTGGTGGCAACAGTTTATGATTCTTCAGATATGTCGAAATGTCAACATTTGCACATA 540
DB 161 LeuTrpSerGluProPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
QY 541 GCTTCAGATCATTTGCCACATTCAGGATTTACTTACAAGACATAAATTTGCTCAGTGCA 600
DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuSerAla 200
QY 601 GAATTTTGGACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
DB 201 GluPheLeuGluGlnHisTyrAspArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
QY 661 GAAATTTATGTGACAAAAGACAGTCACTGAAGCTTCTCGTGTGAACCTACTACTAGATAGA 720
DB 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
QY 721 CACAACCTTCAATTTATGACAAAATATACATCAGTAAACCTGAGAACCTCAAAATTAATGATG 780
DB 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
QY 781 AACTGCTGGCAGACAAAAGTCGCAACATCAGTTTTCAGGCGCTTTCAGCTTTTAAAGTGT 840
DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
QY 841 TTTGTAGCCCATCTTACACAGACGCGCCCATCTCTAGACATCTCTCTCAAGAACCGGCC 900
DB 281 PheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGlnSer
```

Query Match: 89.55% Indels: 9 Gaps: 2
 DB: 2

US-10-089-688-1 (1-1053) x AAH72045 (1-341)

QY 1 ATGCGTTCCTGTTGGAGTCTCAAAATCTCCAGCAGACATTTCTGAAGAACTCTGAAG 60
 DB 1 MetProPheProGlySerHisLysSerProAlaAspLeuValLysAsnLeuLys 20
 QY 61 GAGAGCATGGCTGTTCTGGAAGAGCAGACATTTCTGATAAAAAACGAGAAAGGCTACA 120
 DB 21 GluSerLeuAlaValLeuGluLysGlnAspLeuSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAATTTCCAAAATCTGGTTGCCATGAAGAAATTTCTGATGGCCAAATGAAGAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysThrAsnSerGlyLeuLeu 80
 QY 241 AGCACCTGTGTAGTGTATACAGCTCAATTCAGTCTTTCAGGGCAAAAAAGACGTTGCTCAA 300
 DB 81 GlyThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
 QY 301 ATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTGATATCATC 360
 DB 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrlle 120
 QY 361 TGACCCCAACAGATATTTCTGATGTTATTAAGAGGTATGATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyThrGluSerProGluLeuAla 140
 QY 421 CTAATTTGGAAATATTTAAGAGATGCAATTCAGACATGAACCACTTGCMAAAATCATT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGGTTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCACAAATTTGACATA 540
 DB 161 LeuTrpSerGluProPheTyAspPhePheArgTyThrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAAGACATAAATGCTCAGTGCA 600
 DB 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
 QY 601 GAATTTTGGAAACAGCATATGATGATTTTCTAGTGAATATGAGAAGTTACTTACTTCA 660
 DB 201 GluPheLeuGluGlnHisTyAspArgPhePheSerGluTyThrGluLysLeuLeuHisSer 220
 QY 661 GAAATTTATGACAAAAGACAGTCAGTCAAGCTTCTCGGTGAACTACTACTAGATAGA 720
 DB 221 GluAsnTyThrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
 QY 721 CACAACTTCACAAATATACAAAATATACATCAAGTAAACCTGGAACCTCAAAATTAATGATG 780
 DB 241 HisAsnPheThrIleMetThrLysTyrlleSerLysProGluAsnLeuLysLeuMetMet 260
 QY 781 AACCTGTCGGACAGAAAAGTCGACATCCAGTTCAGTTTGGCCCTTTCAGTTTAAAGTG 840
 DB 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 QY 841 TTTGTAGCCCAATCTTAAACAGCAGCCATCTCTAGACATCTCTCTCAAGAACACGAGCC 900
 DB 281 PheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGlnSer 300
 QY 901 AAACCTCATGAGTTCTCAGCAAGTTTCAAGACGACGAGCGATTTGATGACGAGTTCC 960
 DB 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 QY 961 GTRACGACGACGAAATTCGGGTTCGATTTACGGGTTAAACCGGACGCGGTGGATCAGG 1020
 DB 316 AspGluGlnPheAsnAspGluLysThrTyThrLeuValLysGln-----IleArg 331

QY 1021 GATTTGAAGAGACAGCAGCTCAGCAAGAAAGCT 1050
 DB 332 AspLeuLysArgProAlaGlnGlnAspAla 341

RESULT 6
 Q724X0 PRELIMINARY; PRT; 341 AA.
 AC Q724X0;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-NAR-2004 (TRENBLrel. 26, Last annotation update)
 DE MO25-like protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhou Y., Yu L., Zhao S.Y.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DDJB databases.
 DR EMBL; AF134480; AAP97257.1; --
 DR InterPro: IPR004892; Mo25.
 DR Pfam: PF03204; Mo25; 1
 DR SEQUENCE 341 AA; 39774 MW; D3A60160E78C7A6D CRC64;
 SQ

Alignment Scores:
 Pred. No.: 1.58e-110 Length: 341
 Score: 1590.50 Matches: 323
 Percent Similarity: 92.86% Conservative: 2
 Best Local Similarity: 92.29% Mismatches: 16
 Indels: 9
 Query Match: 88.21% Gaps: 2
 DB:

US-10-089-688-1 (1-1053) x Q724X0 (1-341)

QY 1 ATGCGTTCCTGTTGGAGTCTCAAAATCTCCAGCAGACATTTCTGAAGAACTCTGAAG 60
 DB 1 MetProPheProGlySerHisLysSerProAlaAspLeuValLysAsnLeuLys 20
 QY 61 GAGAGCATGGCTGTTCTGGAAGAGCAGACATTTCTGATAAAAAACGAGAAAGGCTACA 120
 DB 21 GluSerMetAlaValLeuGluLysGlnAspLeuSerAspLysLysAlaGluLysAlaThr 40
 QY 121 GAAGAATTTCCAAAATCTGGTTGCCATGAAGAAATTTCTGATGGCCAAATGAAGAAA 180
 DB 41 GluGluValSerLysAsnLeuValAlaMetLysGluLeuLysGlyThrAsnGluLys 60
 QY 181 GAGCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGCTCCTT 240
 DB 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuLysThrAsnSerGlyLeuLeu 80
 QY 241 AGCACCTGTGTAGTGTATACAGCTCAATTCAGTCTTTCAGGGCAAAAAAGACGTTGCTCAA 300
 DB 81 SerThrLeuValAlaAspLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
 QY 301 ATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTGATATCATC 360
 DB 101 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrlle 120
 QY 361 TGACCCCAACAGATATTTCTGATGTTATTAAGAGGTATGATCTCCAGAAATAGCT 420
 DB 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyThrGluSerProGluLeuAla 140
 QY 421 CTAATTTGGAAATATTTAAGAGATGCAATTCAGACATGAACCACTTGCMAAAATCATT 480
 DB 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIle 160
 QY 481 TTGGTTCGGAACAGTTTATGATTTCTCAGATATGTCGAAATGTCACAAATTTGACATA 540
 DB 161 LeuTrpSerGluProPheTyAspPhePheArgTyThrValGluMetSerThrPheAspIle 180
 QY 541 GCTTCAGATGCAATTTGCCACATTCAGGATTTACTTACAAGACATAAATGCTCAGTGCA 600

Db 191 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuLeuSerAla 200
 Qy 601 GAATTTTGGACAGCAGCATATGATAGATTTTTCAGTGAATATGAGAGTTACTTCAITCA 660
 Db 201 GlupheMetValLeuLeuSerPheArgPhePheSerGluTyrGluLysLeuLeuHisSer 220
 Qy 661 GAAATTTATGTGACAAAAGACAGTCAGTCAAGCTTCTCGGTGAACACTACTACTAGATAGA 720
 Db 221 GluasntyrValThrLysArgGlnSerLeuLysLeuLeuGluLysLeuLeuAspArg 240
 Qy 721 CACAACCTTCAAAATATACAAAATACATCATGTAACCTGAGAACCTCAAAATTAATGATG 780
 Db 241 HisanPheThrIleMetThrLysTyrIleSerLysProGluLysLeuLeuMetMet 260
 Qy 781 AACTGTCTGCGAGACAAAAGTCGCAACATCCAGTTTGGAGCTTTTCACGTTTTTAAGGTG 840
 Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
 Qy 841 TTTGTAGCAATCTTACAAAGCAGCCCATCTCTAGACATCTCTCAAGAACCCAGGCC 900
 Db 281 PheValAlaAsnProAsnLysThrGlnProIleLeuAspIleLeuLeuLysAsnGlnAla 300
 Qy 901 AACTCTAGTAGTTCCTCAGCAAGTTTCAGAACGACAGCAGGAGTGTGATGAGCAGTTC 960
 Db 301 LysLeuLeuGluPheLeuSerLysPheGlnAsnAspArgThrGlu----- 315
 Qy 961 GTACCGACGACGAATTCCTCGGGTCGATTTCACGCTTAAACCGCGAGCGTGGGATCAGG 1020
 Db 316 AspGluGlnPheAsnAspGluLysThrTyrLeuValLysGln-----IleArg 331
 Qy 1021 GATTTGAAGACACAGCTCAGCAGAGAGCT 1050
 Db 332 AspLeuLysArgProAlaGlnGlnGluAla 341

RESULT 7
 Q61QL2 PRELIMINARY; PRT; 341 AA.

AC Q61QL2
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ZGC:86716
 GN Name=ZGC:86716;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McWeeny P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071393; AAH71393.1; -
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR004892; Mo25.
 DR Pfam; PF03204; Mo25; 1.
 SQ SEQUENCE 341 AA; 39792 MW; 9CEB4A665F3C6614 CRC64;

Alignment Scores:
 Pred. No.: 1.45e-106 Length: 341
 Score: 1537.50 Matches: 306
 Percent Similarity: 91.43% Conservative: 14
 Best Local Similarity: 87.43% Mismatches: 21
 Query Match: 85.27% Indels: 9
 DB: 2 Gaps: 1

US-10-089-688-1 (1-1053) x Q61QL2 (1-341)

Qy 1 ATCCGCTTCCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGTGAAGAATCTGAAG 60
 Db 1 MetProPheProPheValLysSerHisLysCysProAlaAspIleValLysAsnLeuLys 20
 Qy 61 GAGACATGCGCTGTTCTGGAAAAGCAGACATTTCTGATAAAAAGCAGAAAAGGTGACA 120
 Db 21 AspAsnMetThrIleLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaSer 40
 Qy 121 GAAGAAGTTTCCAAAATCTGGTTGCCATGAAGAATCTGTATGCGACAAAATGAAAAA 180
 Db 41 GluGluValSerLysSerLeuSerMetLysGluIleLeuTyrGlyThrAsnGluLys 60
 Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTTGCTCAAGAACTCTATATAGTGGGCTCCTT 240
 Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
 Qy 241 AGCACCTGTGTAGCTGATTACAGCTCATTCAGCTTTGAGGGCAAAAAGACGTGGCTCAA 300
 Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysAspValAlaGln 100
 Qy 301 ATTTTCAACAATATTTCTCAGAACACAAATGGTACGAGAACTCTACTGTTGATATCATC 360
 Db 101 IlePheAsnAspIleLeuArgGlnIleGlyThrArgThrProThrValGluTyrLeu 120
 Qy 361 TGCAACCAACAGCAATATTTTGTTCATGTTATTCAGAAAGGTATGAATCTCCAGAAATAGCT 420
 Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProAspIleAla 140
 Qy 421 CTAATTTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAAAAATCAT 480
 Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleThr 160
 Qy 481 TTGTGGTCCGACAGTTTATGATGATTTTCTCAGATATGTCGAATGTCACATTTGACATA 540
 Db 161 LeuCysSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
 Qy 541 GCTTTCAGATGCTATTCGCCACATTCAGAGGATTTACTTACAAGACATATAATTTGCTCAGTGA 600
 Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuSerAla 200
 Qy 601 GAATTTTGGACACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
 Db 201 GluPheLeuGluGlnHisTyrAspPhePheSerGluTyrGluLysLeuLeuHisSer 220
 Qy 661 GAAAAATATGTGACAAAAGACAGTCAGTCAAGCTTCTCGGTGAACACTACTACTAGATAGA 720
 Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGluLysLeuLeuAspArg 240
 Qy 721 CACAACCTTCACAATTAATGACAAATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG 780
 Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
 Qy 781 AACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGGAGCTTTTCACGTTTTTAAGGTG 840

```

Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTGTAGCCAACTCTAACAAGACGACGCCATCTAGACATCTCTCAAGAACACGAGCC 900
Db 281 PheValAlaAsnProAsnLysThrProProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAATCTATAGAGTTCTCCAGCAAGTTTCAGAACGACGAGCGGATTGTATGACAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgAlaGluAspGluGlnPheSer 320
Qy 961 GTACCGACGACGAATCCCGGGTCGATTTCAGCGCTTAAACCGCGGACGCGTGGGATCAGG 1020
Db 321 AspGluLysThrTyr-----LeuIleLysGlnIleArg 331

Qy 1021 GATTGGAAGACAGCAGTCTAGCAAGAAGCT 1050
Db 332 AspLeuLysArgProThrProGlnAspAla 341

RESULT 8
AAH71393
ID AAH71393 PRELIMINARY; PRT; 341 AA.
AC AAH71393;
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton K., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalius D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC071393; AAH71393.1; -.
SQ Hypothetical protein.
SQ SEQUENCE 341 AA; 39792 MW; 9CEE4A665F3C6614 CRC64;

```

```

Alignment Scores:
Pred. No.: 1,45e-106 Length: 341
Score: 1537.50 Matches: 306
Percent Similarity: 91.43% Conservative: 14
Best Local Similarity: 87.43% Mismatches: 21
Query Match: 85.27% Indels: 9
DB: 2 Gaps: 1

```

```

US-10-089-688-1 (1-1053) x AAH71393 (1-341)
Qy 1 ATGCCGTTCCCGTTGGGAAGTCTCAAAATCTCCAGCAGACATCTGTAAGAAATCTGAAG 60
Db 1 MetProPheProPheValLysSerHisLysCysProAlaAspIleValLysAsnLeuLys 20
Qy 61 GAGCAGCATGGCTGTTCTGGAAGAACAGACATTTCTGATAAAAAAGCAGAAAGCGTACA 120
Db 21 AspAsnMetThrIleLeuGluLysGlnAspIleSerAspLysLysAlaGluLysAlaSer 40
Qy 121 GAAGAAGTTTCCAAAATCTGGTGGCCATCAAGAAGAAATCTGTATGGCCAAATGAAGAAA 180
Db 41 GluGluValSerLysSerLeuLeuSerMetLysGluIleLeuTyrGlyThrAsnGluLys 60
Qy 181 GAGCCTCAGACAGACAGTAGCTCAACTTGTCTCAAGAACTCTATATAGTGGGCTCCTT 240
Db 61 GluProGlnThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 80
Qy 241 AGCACCTCGTAGCTGATTACAGCTCATGCTTTGAGGGCAAAAAAGACGTGGCTCAA 300
Db 81 SerThrLeuValAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValAlaGln 100
Qy 301 ATTTTCAACAATATCTCAGACAGACAAATTTGGTACGAGACTCCTACTGTTGAATACATC 360
Db 101 IlePheAsnAspIleLeuArgArgGlnIleGlyThrArgThrProThrValGluTyrLeu 120
Qy 361 TGCACCCCAACAGAAATATTTGTTTATGTTTATTCGAAAGGGTATGAATCTCCAGAAATAGCT 420
Db 121 CysThrGlnGlnAsnIleLeuPheMetLeuLeuLysGlyTyrGluSerProAspIleAla 140
Qy 421 CTAATTTGTGAATAATGTTTAAAGAAATGTCATGACATGACATGACCATTTGCAAAAATCAT 480
Db 141 LeuAsnCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleThr 160
Qy 481 TTGTGGTCGGAACAGATTTTATGATTTCTCAGATATGTCGAAATGTCAACATTTGACATA 540
Db 161 LeuCysSerGluGlnPheTyrAspPhePheArgTyrValGluMetSerThrPheAspIle 180
Qy 541 GCTTCAGATCATTTGCCACATCAAGGATTTACTTACAGACATATAATTTGTCAGTGA 600
Db 181 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuSerAla 200
Qy 601 GAATTTTGAACAGCAGCATATGATATGATATTTTTCAGTGAATATGAGAAGTACTTCAATCA 660
Db 201 GluPheLeuGluGlnHisTyrAspArgPheSerGluTyrGluLysLeuLeuHisSer 220
Qy 661 GAAATTTATGTCACAAAAGACAGTCACTCAAGCTTCTCGTGAACCTACTACTAGATAGA 720
Db 221 GluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 240
Qy 721 CACAATTTCACAATTTATGACAAAATATACATCAGTAAACCTCAGAACCTCAAAATTAATGAT 780
Db 241 HisAsnPheThrIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 260
Qy 781 AACTGCTGCGAGACAAAAGTCGCAACATCCAGTTTGGGCTTTACGTTTTTAAGGTG 840
Db 261 AsnLeuLeuArgAspLysSerArgAsnIleGlnPheGluAlaPheHisValPheLysVal 280
Qy 841 TTTGTAGCCAACTCTAACAAGACGACGCCATCTAGACATCTCTCAAGAACACGAGCC 900
Db 281 PheValAlaAsnProAsnLysThrProProIleLeuAspIleLeuLeuLysAsnGlnThr 300
Qy 901 AAATCTATAGAGTTCTCCAGCAAGTTTCAGAACGACGAGCGGATTGTATGAGCAGTTCC 960
Db 301 LysLeuIleGluPheLeuSerLysPheGlnAsnAspArgAlaGluAspGluGlnPheSer 320
Qy 961 GTACCGACGACGAATCCCGGGTCGATTTCAGCGCTTAAACCGCGGACGCGTGGGATCAGG 1020
Db 321 AspGluLysThrTyr-----LeuIleLysGlnIleArg 331
Qy 1021 GATTGGAAGACAGCAGTCTAGCAAGAAGCT 1050
Db 332 AspLeuLysArgProThrProGlnAspAla 341

```

```

RESULT 9
Q6PAB4 PRELIMINARY; PRT; 337 AA.
AC Q6PAB4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC68674 protein.
GN Name=MGC68674;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_taxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RC MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC060384; AAH60384.1; --
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; MO25.
DR Pfam; PF03204; MO25; 1.
SQ SEQUENCE 337 AA; 39183 MW; CCAB68F5D8569B57 CRC64;

Alignment Scores:
Pred. No.: 9.41e-92 Length: 337
Score: 1339.50 Matches: 268
Percent Similarity: 87.83% Conservative: 35
Best Local Similarity: 77.68% Mismatches: 29
Query Match: 74.29% Indels: 13
DB: 2 Gaps: 4

US-10-089-688-1 (1-1053) x Q6PAB4 (1-337)

QY 1 ATGCGGTTCCGGTTGGGAAGTCTCACAAATCTCCAGCAGACATGTGGAAGATCTCAAG 60
DB 1 MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysThrLeuLys 22

```

```

QY 61 GAGAGCATGGCTGTTCTGGAAAAAGCAAGACATCTTCTGATAAAAAAGCAAGAGCTACA 120
DB 23 AspAsnMetaLaLeuLeuGluArgGln-----AspLysLysThrGluLysAlaSer 39
QY 121 GAAGAAGTTTCCAAAAATCTGGTTCATGAAGAAATCTCTGATGGCACAATGAAAAA 180
DB 40 GluGluValSerLysSerLeuGlnAlaThrLysGluIleLeuCysGlyThrGlyAspLys 59
QY 181 GAGCCTCAGACAGAGCAGTAGCTCAACTCTCTCAAGAACTCTATAATAGTGGGCTCCTT 240
DB 60 GluProGlnThrGluThrValAlaGlnLeuAlaGlnGluLeuTyrAsnSerGlyLeuLeu 79
QY 241 AGCACCTGTGTAGCTGATTACAGCTCATGACCTTTGAGGGCAAAAAAGAGCTGGCTCA 300
DB 80 ValThrLeuIleAlaAsnLeuHisLeuIleAspPheGluGlyLysLysAspValSerGln 99
QY 301 ATTTTCAACAATATCTCAGAGCAAAATGGTTCAGAGAACTCTCTACTGTTGAATACATC 360
DB 100 IlePheAsnAsnIleLeuArgGlnIleGlyThrArgSerProThrValGluTyrIle 119
QY 361 TGCAACCAACAGAAATATTTGTTTCATGTTATTAAGAGGTATGAATCTCCAGAAATAGCT 420
DB 120 SerSerHisGlnHisIleLeuPheIleLeuLeuLysGlyTyrGluSerProGlnValAla 139
QY 421 CTAAATTTGGAATAATGTTAAGAGATGCATCAGACATGAACCACTTGCAGAAATCAT 480
DB 140 LeuHisCysGlyIleMetLeuArgGluCysValArgHisGluProLeuAlaLysValIle 159
QY 481 TTGTGGTCGACAGACAGTTTATGATTTCTTCAGATATCTCGAATATGTCACATTTGACATA 540
DB 160 LeuTyrSerGluGlnPheGlyAspPhePheLysTyrValGluMetSerThrPheAspIle 179
QY 541 GCTTCAGATGCATTGTCACATCAAGAGTTTACTTACAAGACATAAATTTGCTCAGTGCA 600
DB 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysLeuMetValAla 199
QY 601 GAATTTTGAACAGCAGCATATGATAGATTTTTCAGTGAATATGAGAAAGTTACTTCAATCA 660
DB 200 GluPheLeuGluGlnAsnTyrAspArgIlePheAsnAspTyrGluLysLeuLeuHisSer 219
QY 661 GAAATTTATGTGACAAAAGACAGTCATCAAGCTTCTCGTGAACCTACTACTAGATAGA 720
DB 220 GluAsnTyrValThrLysArgGlnSerLeuLysLeuGlyGluLeuIleLeuAspArg 239
QY 721 CACAACCTTCAAAATATGACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATCATG 780
DB 240 HisAsnPheSerIleMetThrLysTyrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACCTGCTCGAGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACTGTTTAAAGGTG 840
DB 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCCCAATCTTAACAGAGCGCAGCCCATCTCTAGACATCTCTCCAGAACCCAGGCC 900
DB 280 PheValAlaAsnProAsnLysThrGlnProIleValAspIleLeuLeuLysAsnGlnThr 299
QY 901 AAATCATAGAGTCTCTCAGCAAGTTTTCAGAACAGCAGGACGAGTGTATGAGCAGTTCC 960
DB 300 LysLeuIleAspPheLeuSerSerPheGlnLysAspArgThrAsp-----Asp 315
QY 961 GTACCGCAGCAAGATTCGCCGGTCGATTTACCGGTTAAACCGCGAGCGGTGGATCAGG 1020
DB 316 GluGlnPheThrAspGluLysAsnTyrLeu-----IleLysGlnIleArg 330
QY 1021 GATTTCAGAGACCA 1035
DB 331 AspLeuLysLysPro 335

```

```

RESULT 10
AAH60384
ID AAH60384 PRELIMINARY; PRT; 337 AA.
AC AAH60384;
DT 02-MAR-2004 (TReMBLrel. 27, Created)

```


or send an email to license@isb-sib.ch).

Alignment Scores:	
Pred. No.:	1,31e-88
Score:	1297.50
Percent Similarity:	85.26%
Best Local Similarity:	76.01%
Query Match:	71.96%
DB:	1
Length:	337
Matches:	263
Conservative:	32
Mismatches:	38
Indels:	13
Gaps:	3

QY	1	ATGCCGTTCCGTTTGGGAAGTCTCAAAATCTCCAGCAGACATTTGGAAGAATCTGAAG	60
Db	4	MetProLeu---PheSerLysSerHisLysAsnProAlaGluIleValLysIleLeuLys	22
QY	61	GAGAGCATGCTGTTCTGGAAGAACCAAGACATTTCTGATAAAAGCAGAAAAGCGCTACA	120
Db	23	AspAsnLeuAlaIleLeuGluLysGln-----AspLysLysThrAspLysAlaSer	39
QY	121	GAGAAATTTCCAAAAATCTGGTTGGCCATGAAAGAAATTCGTATGGCACAATAGAAAA	180
Db	40	GluGluValSerLysSerLeuGlnAlaMetLysGluIleLeuCysGlyThrAsnGluLys	59
QY	181	GAGCCTCAGACAGAAGCAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGGCTCCT	240
Db	60	GluProProThrGluAlaValAlaGlnLeuAlaGlnGluLeuTyrSerSerGlyLeuLeu	79
QY	241	AGCACCTGGTAGCTGATTTACAGCTCATTTAGCTTTGAGGGCAAAAAAGACGTGGCTCAA	300
Db	80	ValThrLeuIleAlaAspLeuGlnLeuIleAspPheGluGlyLysLysAspValThrGln	99
QY	301	ATTTTCAACAATATTCTCAGAAGACAAATTGGTACGAGAACTCTACTGTTTGAATACATC	360
Db	100	IlePheAsnAsnIleLeuArgArgGlnIleGlyThrArgSerProThrValGluTyrIle	119
QY	361	TGCACCCACAGAAATATTTTGTTCATGTTATTGAAGGGTATGAATCTCCAGAAATAGCT	420
Db	120	SerAlaHisProHisIleLeuPheMetLeuLeuLysGlyTyrGluAlaProGlnIleAla	139
QY	421	CTAAATGTGGAATAATGTTTAAGAGAAATGCATCAGACATCAACACATTCGCAAAATCAT	480
Db	140	LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleIle	159
QY	481	TTGTGFTGGAACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCACAAATTTGACATA	540
Db	160	LeuPheSerAsnGlnPheArgAspPhePheLysTyrValGluLeuSerThrPheAspIle	179
QY	541	GCATTACAGTCATTTGCCACATTCAGAGATTTACTTACAAGACATAAATGCTCAGTGCA	600
Db	180	AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla	199
QY	601	GAAATTTTGGAAACGACATTAATGATAGATATTTTTCAGTGAATATGAGAAGTTACTTCA	660

```

QY 181 GAGCTCAGACAGAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTT 240
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuTySerSerGlyLeuLeu 79
QY 241 AGCACCTGGTAGCTGATTACAGCTCACTGCTTTCAGGGCAAAAGAGCTGGCTCAA 300
Db 80 ValThrLeuLeuAlaAspLeuGlnLeuLeuAlaAspPheGluGlyLysAspValThrGln 99
QY 301 ATTTTCAACATATTCTCAGNAGACAAATTTGGTACGAGAACTCTACTGTTGATACATC 360
Db 100 IlePheAsnAniLeuLeuArgGlnIleGlyThrArgCysProThrAlaGluTyrlle 119
QY 361 TGCACCCCAACAGATATTTTGTTCATGTTTAAAGAGGTATGAATCTCCAGAAATAGCT 420
Db 120 SerSerHisProHisIleLeuValMetLeuLeuLysGlyTyThrGluAlaProGlnIleAla 139
QY 421 CTAATTTGTGAATAATTTAAAGAGATGTCATCAGACATGAACCACTTGCAGAAATCATT 480
Db 140 LeuArgCysGlyIleMetLeuArgGluCysIleArgHisGluProLeuAlaLysIleile 159
QY 481 TTGTGGTCGGACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCACATTTGACATA 540
Db 160 LeuPheSerAsnGlnPheArgAspPhePheLysTyThrValGluLeuSerThrPheAspIle 179
QY 541 GCTTCAGATGCTATTTGCCACATTCAGAGATTCTTCAAGACATATAATGCTCAGTGCA 600
Db 180 AlaSerAspAlaPheAlaThrPheLysAspLeuLeuThrArgHisLysValLeuValAla 199
QY 601 GAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCAATCA 660
Db 200 AspPheLeuGluGlnAsnTyThrIlePheGluAspTyThrGluLysLeuLeuGlnSer 219
QY 661 GAAATATGTCAGACAAAAGACAGTCACTGAGCTTCTCGGTGAAGTCACTACTAGATAGA 720
Db 220 GluAsnTyThrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAspArg 239
QY 721 CACAACTTCACAATATGACAAAATACATCATGATAAACCCTGAGAACCTCAAAATTAATGATG 780
Db 240 HisAsnPheThrIleMetThrLysTyThrIleSerLysProGluAsnLeuLysLeuMetMet 259
QY 781 AACCTGCTGCAGACAAAAGTCGCAACATCAGATTGAGGCTTTTTCAGCTTTTAAAGTG 840
Db 260 AsnLeuLeuArgAspLysSerProAsnIleGlnPheGluAlaPheHisValPheLysVal 279
QY 841 TTTGTAGCAATCTTAACAGACGAGCCCATCTCTAGACATCTCTCTCAAGAACAGGCC 900
Db 280 PheValAlaSerProHisLysThrGlnProIleValGluIleLeuLeuLysAsnGlnPro 299
QY 901 AAATCTAGAGTTCTCTCAGCAAGTTTTCAGAACGACGAGCGGATTTGATGAGCAGTTCC 960
Db 300 LysLeuIleGluPheLeuSerSerPheGlnLysGluArgThrAsp----- 314
QY 961 GTACCGACGAGAAATTCCTCGGGTGCATTTCACGCGTTAAACCGCGACGCGTGGATCAGG 1020
Db 315 -----AspGluGlnPheAlaAspGluLysAsnTyThrLeuIleLysGlnIleArg 330
QY 1021 GATTGAGAGACAGCAGCT 1038
Db 331 AspLeuLysLysAlaAla 336

```

RESULT 14

```

QY 181 GAGCTCAGACAGAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTT 60
Db 60 GluProProThrGluAlaValAlaGlnLeuAlaGlnLeuTySerSerGlyLeuLeu 20
QY 61 GAGAGCATGGCTGTTCTGGAAAAGCAAGACATTTCTCAT---AAAAAGCAGAAAAGGCT 117
Db 21 GluAsnValAlaTyThrLeuGluLysLeuGluSerSerGluSerLysLysCysGluLysVal 40
QY 118 ACAGACAGAGTTTCCAAAAATCTGTTGCCATGAAAGAAATTTCTGTATGGCACAATCAA 177
Db 41 AlaGluGluValSerLysAsnLeuSerSerLeuLysGluValLeuLysGlyThrGlyAsp 60
QY 178 AAAGAGCTCAGACAGACAGTAGCTCACTCACTTCTCAAGAACTCTATATAGTGGGCTC 237
Db 61 LysGluProGlnThrGluAlaValAlaGlnLeuAlaGlnLeuTyThrAsnThrAsnLeu 80
QY 238 CTTAGCACCTGGTAGCTGATTACAGCTCATTCTAGGTCATTTAGGGCAAAAAGAGCTGGCT 297
Db 81 LeuLeuSerLeuIleAlaAsnLeuGlnArgIleAspPheGluGlyLysLysAspValVal 100
QY 298 CAAATTTTCAACAATATTCTCAGAGAGCAAAATTTGGTACGAGAACTCTCTACTGTTGAATAC 357
Db 101 HisLeuPheSerAsnIleValArgArgGlnIleGlyAlaArgThrProThrValGluTyThr 120
QY 358 ATCTGACCCCAACAGATATTTTGTTCATGTTTATTTGAAGGTTATGATCTCCAGAAATA 417
Db 358 ATCTGACCCCAACAGATATTTTGTTCATGTTTATTTGAAGGTTATGATCTCCAGAAATA 417

```

```

OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC04172; AAH4172.1; -.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
SQ SEQUENCE 343 AA; 39820 MW; F10450DA0446268A CRC64;

```

```

Alignment Scores:
Pred. No.: 7,97e-88 Length: 343
Score: 1287.00 Matches: 253
Percent Similarity: 83.33% Conservative: 42
Best Local Similarity: 71.47% Mismatches: 41
Query Match: 71.38% Indels: 18
DB: 2 Gaps: 3

```

US-10-089-688-1 (1-1053) x Q803V8 (1-343)

```

Db 121 IleSerSerHisSerGlnIleLeuPheMetLeuLeuLysGlyTyrGluThrSerGluVal 140
Qy 418 GCTCTAAATTTGGAATAATTTAAGAGATGCAATGACATGAAACCACTTGCAGAAATC 477
Db 141 AlaLeuAsnGlyMetMetLeuArgGlyCysLeuArgHisAspProLeuAlaArgile 160
Qy 478 ATTTTGTGTCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCACAAATTCAC 537
Db 161 ValLeuPheSerGluAspPheTyrCysPhePheArgTyrValGluMetSerThrPheAsp 180
Qy 538 ATAGCTTCAGATGATTTGCCACATTTCAAGATTTTACTTACAGACATAAATTCCTCAGT 597
Db 181 IleAlaSerAspAlaPheAlaSerPheLysAspLeuLeuThrArgHisLysIleMetCys 200
Qy 598 GCAGAAATTTTGGAAACAGCATTTATGATATTTTTCAGTGAATGAGAAATTTACTTCAT 657
Db 201 AlaAspPheLeuGluThrAsnTyrAspArgValPheThrGluTyrGluLysLeuLeuHis 220
Qy 658 TCAGAAATTTATGTCAGAAAGACAGTCACAGCTTCTCGTGAACCTTACTACTAGAT 717
Db 221 SerGluAsnTyrValThrLysArgGlnSerLeuLysLeuLeuGlyGluLeuLeuAsp 240
Qy 718 AGACAACTTCACAAATTTATGACAAAATACATCAGTAAACCTCGAGAACCTCAAAATTAATG 777
Db 241 ArgHisAsnPheThrValAlaThrLysTyrIleSerArgAlaGluAsnLeuLysLeuMet 260
Qy 778 ATGAACCTGTCGAGACAAAAGTCGCAACATCCAGTTTGAAGCTTTCACGTTTAAAG 837
Db 261 MetAsnMetLeuArgAspAsnSerArgAsnIleGlnPheGluAlaPheHisValPheLys 280
Qy 838 GTGTTTTCAGCAATCTTACAAACGACGACCCATCTAGACATCTCCTCAGAACACAG 897
Db 281 ValPheValAlaAsnProAsnLysThrGlnProValLeuAspIleLeuLeuLysAsnGln 300
Qy 945 GCCAAATCTACAGATTTCTTCAGCAAGTTTTCAGAACGACGACGACGAT----- 945
Db 301 SerLysLeuValGluPheLeuSerHisPheGlnThrAspArgSerGluAspGluGlnPhe 320
Qy 946 TGTATGAGCATGTCCTGACGACGACGAAATTCCTCGGTTCGATTTACCGTTAAACCCGCG 1005
Db 321 CysAspGluLysAsnTyr----- 327
Qy 1006 ACGGTGGGATCAGGATTTTCAGAGACCACTCAGCAAGNA 1047
Db 328 IleLysGlnIleArgAspLeuLysArgProAlaProAlaGlu 341

RESULT 15
Q7PF07 PRELIMINARY; PRT; 332 AA.
AC Q7PF07;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000023148 (Fragment).
GN Name=ENSANGG0000021597;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008804; EAA45510.1; -.
DR InterPro; IPR004892; Mo25.
DR Pfam; PF03204; Mo25; 1.
FT NON TER 332 332
SQ SEQUENCE 332 AA; 38665 MW; E315A6819F3BDA67 CRC64;

```

```

Alignment Scores: 1.38e-78 Length: 332
Pred. No.: 1163.50 Matches: 226
Score: 86.08% Conservative: 46
Percent Similarity: 71.52% Mismatches: 39
Best Local Similarity: 64.53% Indels: 5
Query Match: 2 Gaps: 3
DB: 2

US-10-089-688-1 (1-1053) x Q7PF07 (1-332)

Qy 1 ATCCGCTTCCCGTTGGGAAGTCTCACAAATCTCCAGCAGACATTTGTGAAGAAATCTGAAG 60
Db 1 MetProLeu---PheGlyLysSerGlnLysSerProGlnGluLeuValLysSerLeuLys 19
Qy 61 GAGACATGCTGCTGTTTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGCTACA 120
Db 20 GluAlaValAlaAsnAlaLeuGluArg-----GlyAspLysLysAlaGluLysAlaGln 36
Qy 121 GAAGAAGTTTCCAAAATCTGTTGCCATGAAGAATAATCTGTATGCGCACAATGAAAAA 180
Db 37 GluAspValSerLysAsnLeuValSerIleLysAsnMetLeuCysGlyThrAlaAspAla 56
Qy 181 GAGCCTCAGACAGAA---GCAGTAGCTCAACTTGTCTCAAGAACTCTATAATAGTGGGCTC 237
Db 57 GluProGlnThrGluLysValIleSerGlnLeuAlaHisGluLeuTyrSerThrAspLeu 76
Qy 238 CTTAGCACCTGCTGATGATTTACAGCTCATTGACCTTTGAGGGCAAAAAGACGCTGCT 297
Db 77 LeuLeuLeuLeuIleGlnAsnLeuAsnArgIleAspPheGluGlyLysLysAspValAla 96
Qy 298 CAATTTTCAACATATTTCTCAGAAAGCAAAATTTGGTACGAGAACTCTACTGTTGAATAC 357
Db 97 HisIlePheAsnAsnValIleArgArgGlnIleGlyThrArgLeuProThrValGluTyr 116
Qy 358 ATCTGCACCCCAACAGAAATATTTGTTTCATGCTTATTGAAGGCTATGAATCTCCAGAAATA 417
Db 117 IleCysThrLysProGluLysLeuPheThrLeuMetAlaGlyTyrGluHisGlnGluLe 136
Qy 418 GCTCTAAATTTGGAATAATTTAAGAGAAATGCAATCAGACATGAACCACTTGCAGAAATC 477
Db 137 AlaLeuAsnGlyIleMetLeuArgGluCysAlaArgHisGluAlaLeuAlaLysIle 156
Qy 478 ATTTTGTGTCGGAACAGATTTTATGATTTCTCAGATATGTCGAAATGTCACAAATTTGAC 537
Db 157 MetLeuHisSerGluGluPhePheAsnPhePheArgTyrValGluValSerThrPheAsp 176
Qy 538 ATAGCTTCAGATGCTTGGCAGATTTCAAGGATTTACTTACAAGACATAAATTCCTCAGT 597
Db 177 IleAlaSerAspAlaPheSerThrPheLysGluLeuLeuThrArgHisLysLeuLeuSer 196
Qy 598 GCAGAAATTTTGGAAACAGCAATTTATGATAGATTTTTCAGTGAATATGAGAAATTTACTCAT 657
Db 197 AlaGluPheLeuGluGlnAsnTyrAspLysValPheSerArgTyrGluAlaLeuLeuAsn 216
Qy 658 TCAGAAATTTATGTCAGAAAAGACAGTCACTGAGCTTCTCGTGAACCTACTACTAGAT 717
Db 217 SerGluAsnTyrValThrArgGlnSerLeuLysLeuLeuGlyGluLeuLeuLeuAsp 236
Qy 718 AGACAACTTTCACAAATTTATGACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATG 777
Db 237 ArgHisAsnPheThrValMetThrLysTyrIleSerAsnProAsnLeuLeuLeuMet 256
Qy 778 ATGAACCTGTCGAGACAAAAGTCGCAACATCCAGTTTTCAGGGCTTTTCAGTTTTTAAG 837
Db 257 MetAsnMetLeuLysGluThrSerArgAsnIleGlnPheGluAlaPheHisValPheLys 276
Qy 838 GTGTTTGTAGCCAAATCTTACAAAGACGACCCCATCTAGACATCTCCTCCTCAAGAACAG 897
Db 277 ValPheValAlaAsnProAsnLysProLysProIleMetAspIleLeuLeuLeuArgAsnGln 296
Qy 898 GCCAAATCTACAGATTTCTCAGCAAGTTTTCAGAACGACGACGACGAT 945
Db 297 GluLysLeuValAspPheLeuThrArgPheHisThrAspArgSerGlu 312

```

Search completed: November 10, 2004, 20:51:01
Job time : 247.5 secs

• • • • •


```
QY 651 ----- 651
Db 2993 CysSerThrValSerValAspSerAlaAsnProProAspLeuAenLysMetTrpAsnGlu 3012
QY 652 ---CTTCATTAGAAATATGTGACAAA-----AGACAGTCACCTCAAGCTTCTCGGT 702
Db 3013 ProPheTyrGlnGluThrTyrLeuProTyrMetIleArgSerLysLeuLysLeuLeu 3032
QY 703 GAATCTACTAGTAGACACAACTTCACAAATATGACAAATATACATCAGTAACCTGAG 762
Db 3033 GlnGlyGluGlyAspGlnSerLeuLeuThrPheIleAspGluAlaValSerLys---Glu 3051
QY 763 AACCTCAAAATTAAGTGAACCTG-----CTG 789
Db 3052 LeuGlnLysValLeuValGlnLeuHisTyrSerGlnGlnLeuSerLeuLysLeuLysLeu 3071
QY 790 CGACACAAAATGTCACACATCCAGTTTGAGGCTTTTCACGTTTTTAAGGTGTTGTAGCC 849
Db 3072 GlnAspAspValAspArgAlaLysTyrTyrIleGlnAenCysIleArgIlePheMet--- 3090
QY 850 AATCCTAACAGACGCGACCCATCTAGACATCTCTCTCAAGAACCCAG----- 897
Db 3091 -----GlnSerTyrSerSerIleAspValLeuLeuGlnArgSerArgLeuThrLys 3107
QY 898 -----GCCAACTCATAGATTCTCTCAGCAAGTTT 927
Db 3108 LeuGlnSerLeuGlnAlaLeuIleGlnIleGlnPheIleSerPheIleArgLysGln 3127
QY 928 CAGAACGACAGGACGGATTGTATGACGAGTTCCTGACCGACG----- 969
Db 3128 GlyAsn-----LeuSerSerGlnIleProLeuLysArgLeuLeuLysThr 3142
QY 970 ---ACGAATCCCGGTGCGATTACGGTTAAACCG 1002
Db 3143 TriPThrAsnArgTyrProAspAlaLysMetAspPro 3154

RESULT 3
US-10-732-923-3091
; Sequence 3091, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3091
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-732-923-3091

Alignment Scores:
Pred. No.: 0.0194 Length: 793
Score: 105.50 Matches: 75
Percent Similarity: 37.15% Conservative: 45
Best Local Similarity: 23.22% Mismatches: 132
Query Match: 5.85% Indels: 71
DB: 6 Gaps: 15

US-10-089-688-1 (1-1053) x US-10-732-923-3091 (1-793)
QY 172 AATGAAAAGAGCTCAGACAGACAGTAGCTCAACTTCTCAGAACTCTATAAGT 231
Db 139 AspGlnProSerProSerSerHisArgValSerAspPheSerAspGlu---TyrSerPro 157
QY 232 GGGCTCTTACGACCCCTGGTAGCTGATTACAGCTCATTCATTCTTGGGGCAAAAAGAC 291
Db 158 SerLysValArgIleLeuValAlaAsp-----AspAspAsp 169
QY 292 GTG---GCTCAAAATTTCAACAATATTTCTCAGAAGA-----CAAATT 330
```

```
Db 170 ValGlnArgLysIleLeuAenAsnLeuLysPheHisTyrAsnValThrLeuVal 189
QY 331 GGTACGAGAACTCTCTACTGTTGAATACATCTGCACCCCAACAGAAATATTTGTTTCATGTTA 390
Db 190 ProAenGlyGluIleAlaTrpGluTyrIleAenLysGlyGlnGlnLysTyrAspLeuVal 209
QY 391 TTGAAAGGGTATGATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTTAAGAAATGC 450
Db 210 LeuThrAspValMetMetProHisIleThr-----GlyPheAspLeuLeuGlnArg 226
QY 451 ATCAGACATCAACCATTCGAAAA-----ATCATTTTGTGGTCGGAACAGTTTAT 501
Db 227 IleAsnAspHisProValHisArgHisIleProValIleLeuMetSerGlyThrAlaVal 246
QY 502 GATTTCCTCAGATATGCGAAATGTCACATATTTGCATAGCTTCAGATGCTATTTGCCACA 561
Db 247 Asp---TyrLysTyrAlaAsn---AspThrIleLysIleGlyGlyGlnAspPheLeuThr 264
QY 562 TTCAGGATTTACTTACACAGACATAATCTCAGTCAGCCAGAAATTTTGGACACAGCATAT 621
Db 265 -----LysProIleAlaLysGluLeuLysLysLysIle 276
QY 622 GAT-----AGATTTTTCAGTGAATATGAACTTACTT 654
Db 277 AspThrValLeuGlnSerIleTrpGlnArgLysGluGluGluTyrLysAlaThrLeu 296
QY 655 CATTCAGAAATATGTCACAAAAGACAGCTCACTGAAGCTTCTCGTGAACTACTACTACTA 714
Db 297 AlaGlnGlu-----ArgGluLysGlyAsnLysLeuAlaLysGluMetGluLeu 312
QY 715 GATAGACACAACTTCACAAATATGACAAAATATACATCATGTAACCTGAGAACCTC----- 768
Db 313 LysGluHisGluIleGluGluLeuThrLysLysValSerLysMetSerSerIleSerLys 332
QY 769 AATTAATGATGAACCTGCTCGCAGACAAAAGTCGCAACATCCAG----- 813
Db 333 GluAlaMetGluSerProLeuValSerValThrArgAsnIleGluGluLeuLysGln 352
QY 814 -----TTGAGGCTTTTACGCTTTTAAAGGTGTTTGTAGCAATCTCTAAC 858
Db 353 SerSerTrpSerHisTyrGluSer-----GluIleLysGlu 364
QY 859 AAGACGACGCCCATCTCTAGACATCTCTCCTCAAGAACCCAGCCCAAACTCATAGAGTTCTC 918
Db 365 LysLeuSerSerIleLeuLysGluLeuGlySerSerAsnIleTyrArgProSerPheGlu 384
QY 919 AGCAAGTTTCAGAACGACGACGAGCGATTGTATGACGAGTTCCGTA----- 963
Db 385 LysLeuIleLysAsnAspSerValAspProValThrLysSerPheLeuValSerGluPhe 404
QY 964 ---CCGACGACGAATTCCTCGGTCGATTACGCGTTAAACCGCGACGCTGGGATCAGG 1020
Db 405 SerSerThrThrSerArgAsnSerIleProThrPheProGlnThrThrTyrAsnArg 424
QY 1021 GATTGGAAG 1029
Db 425 AspThrLys 427

RESULT 4
US-10-952-698-164
; Sequence 164, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: OF SCREENING FOR MODULATORS OF BLADDER CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/952,698
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 164
```

```
; LENGTH: 1380
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-164

Alignment Scores:
Pred. No.: 0.0347 Length: 1380
Score: 103.50 Matches: 90
Percent Similarity: 39.01% Conservative: 59
Best Local Similarity: 23.56% Mismatches: 136
Query Match: 5.74% Indels: 97
DB: 6 Gaps: 21

US-10-089-688-1 (1-1053) x US-10-952-698-164 (1-1380)
QY 13 TTTGGGAAGTCTCACAATCTCCAGCAGACATGTTGTAAGATCTG-----57
Db 888 PheLysLysGluAsnGluThrLeuLysSerAspLeuAsnLeuMetGluLeuGlu 907
QY 58 -----AAGGAGAGCATGGCTGTTCTGAAAGCAAGACATTTCTGATAAAGACAGAA 111
Db 908 AlaGluLysGluArgAsnAsnLysLeuSerLeuGlnPheGluGluAspLys-----Glu 925
QY 112 AAGCTCAGAGAAGTTCACAAATCTGGTTCCTCCATGAAGAAATCTGTATGGCACA 171
Db 926 AsnSerSerLysGluLeuLysValLeuLysValLeuGluAlaValArgGln-----940
QY 172 AATGAAAAAGAGCCTCAGACAGACAGACAGTACTCAA-----CTTGCTCAGAACTC 222
Db 941 ---GluLysGlnLysGluThrAlaLysCysGluGlnMetAlaLysValGlnLysLeu 959
QY 223 TATAATAGTGGCTCCTTAGCACC-----CTGTAGTCTGATTTACAGCTCATTGACTTT 276
Db 960 GluLysSer---LeuLeuAlaThrGluLysValLysSerLeuGluLysSerArgAsp 978
QY 277 GAGGCGAAAAAGACGTGGCTCAAAATTTCAACAATATT-----CTCAGAAGCAAAAT 330
Db 979 SerAspLysLysValValAlaAspLeuMetAsnGlnLeuGluLeuArgThrSerVal 998
QY 331 GGTACGAGAACTCTACTGTTGATATACATCTGCACCCACAGATATTTGTCATGTTA 390
Db 999 CysGluLysThrGluThrIleAsp-----ThrLeu 1008
QY 391 TTGAAGAGGTATGAATCTCCAGAAATAGCTCTAAATTTGGGAATAATTTAAGAGATGC 450
Db 1009 LysGlnGluLeuLysAspIleAsnCysLysTyrsAsnSerAlaLeuValAspArgGluGlu 1028
QY 451 ATCAGACATGAACCACTTGCAAAAATCATTTTGGTGGGAACAGTTTATGATTTCTTC 510
Db 1029 SerArg-----ValLeuIleLysLysGlnGluValAspIleLeu 1041
QY 511 AGATATGCGAAATGTCACATTTGACATGCTTCAGATGCTATTTGCCACATTCACAGAT 570
Db 1042 AspLeuLysGluThrLeuArgLeuArgIleLeuSerGluAspIleGlu-----ArgAsp 1059
QY 571 TTACTT-----ACAGACATAAATTTGCTCAGTGCAGAAATTTTG 609
Db 1060 MetLeuCysGluAspLeuAlaHiAlaThrGluGlnLeuAsnMetLeuThrGluAlaSer 1079
QY 610 GAACAGCATATATGATATTTTTCAGTGAATATATGAGAAGTTACTTCAATTCAGAAAAATAT 669
Db 1080 LysLysHisSerGlyLeuLeuGlnSerAlaGlnGluGluLeu-----1093
QY 670 GTGCAAAAAGACAGTCTACTAGCTTCTCGGTGAACCTACTACTAGATAGACACAACTTC 729
Db 1094 ---ThrLysLysGluAla-----LeuIleGlnGluLeu-----GlnHis-----1105
QY 730 ACAATTATGACAAAATATACATCAGTAAACCTCAGAACCTCAAAATTAATGATGAACCTGCTG 789
Db 1106 -----LysLeuAsnGlnLysLysGluGluValGluGlnLysLysAsnGluTy 1121
QY 790 CGAGACAAAAGTCGCAACATCCAGTTTGGAGCCTTTTACAGTTTAAAGGTGTTGTAGCC 849
Db 1122 AsnPheLysMetArgGlnLeuGlu-----HisValMetAspSerAlaAlaGlu 1137
QY 850 AATCCTTAAC-----AAGACGCGAGCCC-----ATCCTAGACATCTCTCTCAAG 891
Db 1138 AspProGlnSerProLysThrProHisPheGlnThrHisLeuAlaLysLeuGlu 1157
QY 892 AACGAGGCAAACTCATAGAG-----TTCTCTC 918
Db 1158 ThrGlnGluGlnGluIleGluAspGlyArgAlaSerLysThrSerLeuGluHisLeuVal 1177
QY 919 AGCAAGTTTTCAGAACACGACGACGCGGATTGTATGACGAGTCCGTACCGACGACGATTC 978
Db 1178 ThrLysLeuAsnGluAspArg-----GluValLysAsnAla 1189
QY 979 CGGTCGATTACGCTTAAACCGCGCGCGTGGGATCAGGATTTGAAGACGACGACT 1038
Db 1190 GluIle---LeuArgMetLysGluGlnLeuArgGluMetGluAsnLeuArgLeuGluSer 1208
QY 1039 CAGCAA 1044
Db 1209 GlnGln 1210

RESULT 5
US-10-967-702-245
; Sequence 245, Application US/10967702
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Pauloski, Nicole
; APPLICANT: Taylor, Ian
; APPLICANT: Bigwood, Douglas
; FILE REFERENCE: 5176
; TITLE OF INVENTION: GENE EXPRESSION PROFILES AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/10/967,702
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US 60/508,355
; PRIOR FILING DATE: 2003-10-03
; NUMBER OF SEQ ID NOS: 400
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 245
; LENGTH: 1583
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-967-702-245

Alignment Scores:
Pred. No.: 0.0359 Length: 1583
Score: 103.50 Matches: 84
Percent Similarity: 36.57% Conservative: 59
Best Local Similarity: 21.48% Mismatches: 133
Query Match: 5.74% Indels: 115
DB: 6 Gaps: 19

US-10-089-688-1 (1-1053) x US-10-967-702-245 (1-1583)
QY 19 AAGTCTCACAATCTCCAGCAGACATTTGTAAGAAATCTGAAGGAGACATGGCTGTT---75
Db 43 LysTySerGluAspLysAlaAsnLeuGlnLysGlnLeuGluAlaMetAsnThrGln 62
QY 76 -----CTGGAAGCAACAGACATTTCTGATAAAGACGAGAAAGGCT 117
Db 63 LeuGluLeuSerGluGlnLeuLysPheGlnAsnSerGluAspAsnValLysLysLeu 82
QY 118 ACAGAAGAAGTTTCCAAAAATCTGTTGCCATGAAGAA---ATTCTGTAT-----165
Db 83 GlnGluGluIleGluLysIleArgProGlyPheGluGlnIleLeuTyLeuGlnLys 102
QY 166 -----GGCAAAATGAAGAGAG-----CCTCAGACAGACGAGTAGCT 204
Db 103 GlnLeuAspAlaThrThrAspGluLysLysGluThrValThrGlnLeuGlnAsnIle 122
QY 205 CAATCTGCTCAAGNACTCTATAGTGGCTCTCTAGCACCTGCTGATGATTTACAG 264
Db 123 GluAlaAsnSerGlnHisTyGlnLysAsnIleAsnSerLeuGlnGluLeuGln 142
```

```

QY 265 CTC-----ATTGACTTTGAGGC 282
Db 143 LeuylsalalleHISglnGluVallysGluLeuMetCysGlnIleGluAlaSerAla 162
QY 283 AAAAAAGACGTGGCTCAAAATT-----TTCAACAATATTCTCAGA 321
Db 163 LysGluHISGluAlaGluileAsnLysLeuAsnGluLeuLysGluAsnLeuVallysGln 182
QY 322 AGACAAATTGTCAGAGACTTCCTACTGTTGAATACATCTGCACCAACAGAAATATTG 381
Db 183 CysGluAlaSerGluLysAsnIleGlnLysLysTyrGluCysGluLeuGluAsnLeuArg 202
QY 382 TTCATGTTATTGAAGGGTATGAATCTCCAGAAATAGCTCTAAATTTCTGGNATATCTTA 441
Db 203 LysAlaThrSerAsnAlaAsnGlnAspAsnGlnIle-----CysSerIleLeu 219
QY 442 AGAGATGTCATCAGACATGAACCACTTGCAGAAATCATTTTGTGTCGGAACAGTTTAT 501
Db 220 GlnGlu----- 221
QY 502 GATTCTTCAGATATGCGAAATGTCACATTTGCACATAGCTTCAGATGCAATTTGCCACA 561
Db 222 -----AsnThrPheValGluInValValAsnGluLysVallysHisLeuGluAspThr 239
QY 562 TTCAGAGTTTACTTACAGACATAAATTTGCTAGTCAGCAATTT----- 606
Db 240 LeuLysGluLeuGluSerGlnHisSerIleLeuLysAspGluValThrTyrMetAsnAsn 259
QY 607 -----TTGGAA-----CAGCATTAT--GATAGATTTTTCAGTGAATAT 642
Db 260 LeuLysLeuLysLeuGluMetAspAlaGlnHisIleLysAspGluPhePheHisGluArg 279
QY 643 GAGAGTTACTTTCAT-----TCAGAAATTTATGTGACAAAAAGACAGTCA----- 687
Db 280 GluAspLeuGluPheLysIleAsnGluLeuLeuLeuAlaLysGluGlnGlnGlyCysVal 299
QY 688 -----CTGAAGTTCTCGGTGAACATA-----CTACTAGAT 717
Db 300 IleGluLysLeuLysSerGluLeuAlaGlyLeuAsnLysGlnPheCysTyrThrValGlu 319
QY 718 AGACACAATCTCAAAATTTATGACA-----AAATACATC----- 750
Db 320 GlnHisasnArgGluValGlnSerLeuLysGluGlnHisGlnLysGluLysSerGluLeu 339
QY 751 -----AGTAAACCTCGAACCTCAAAATTAATGATGAACCTG--- 786
Db 340 AsnGluThrPheLeuSerAspSerGluLysGluLysLeuThrLeuMetPheGluIleGln 359
QY 787 ---CTGCGACACAAAAGTCGACATCCAGTTTGAGGCCTTTTCACGTTTTTAAGGTGTTT 843
Db 360 GlyLeuLysGluGlnCysGluAsnLeuGlnGlnGluLys-----GlnGluAla 375
QY 844 GTAGCCAACTTAACAGACGCGCCCTCTAGACATCTCTCAGAACCCAGGCCAAA 903
Db 376 IleLeuAsnTyrGluSerLeuArgGluLeuMetGluLeu-----GlnThrGlu 392
QY 904 CTCATAGAGTTCTCAGCAAGTTTTCAGAACGAC 936
Db 393 LeuGlyGluSerAlaGlyLysIleSerGlnGlu 403

```

RESULT 6

```

US-10-732-923-8275
; Sequence 8275, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES.
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10732,923
; PRIOR FILING DATE: 2003-12-10
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149

```

```

; SEQ ID NO 8275
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Mycoplasma genitalium
US-10-732-923-8275

Alignment Scores:
Pred. No.: 0.0353 Length: 600
Score: 102.50 Matches: 73
Percent Similarity: 38.80% Conservative: 69
Best Local Similarity: 19.95% Mismatches: 151
Query Match: 5.68% Indels: 73
DB: 6 Gaps: 14

```

US-10-089-688-1 (1-1053) x US-10-732-923-8275 (1-600)

```

QY 88 GACATTTCATATAAAAGACAGAA----- 111
Db 66 AspleuasnaspGluasnAlaLysAlaLeuIleSerPheTrpSerGlnIleAspLeuArg 85
QY 112 AAGGTACAGAGAAGTTTCCAAAAATCTGGTTGCCATGAAAGAAATTTCTGTATGCG--- 168
Db 86 AsnAlaThrCysAspAlaAsnAsnLeuLeuAlaLysTyrHisLeuPheGluVal 105
QY 169 -----ACAAATCAAAAAGACCTCAGACAGAAGCA 198
Db 106 ValAspThrIleSerLysTyrThrAspValThrAsnLeuLeuAspGluLysThrAlaLeu 125
QY 199 GTAGCTCAACTCTCTCAAGAA---CTCTATAATAGTGGGCTCTTAGCACCTGGTAGCT 255
Db 126 GlnLeuLysIleAlaSerGluAsnGlnPheHisSerLeuAspAsnAsnGlnIleAsn 145
QY 256 GATTTCAGCTCATTTGAGGCGCAAAAAGACGTGGCTCAAAATT----- 303
Db 146 AsnLeuArgLysTyrGlnLysGluGlyValLysTrpIleArgAlaLeuGluAspAsnGln 165
QY 304 TTCACAAATATTTCTCAGAGACAAATTTGTACAGAACTCTTACTGTTGAATACATCTGC 363
Db 166 PheGlyGlyIleLeuAlaAspGluMetGly-----LeuGly 177
QY 364 ACCCAACAGAAATATTTGTTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCTA 423
Db 178 LysThrAlaGlnValIlePheAlaMetLeuAspSerTyrGlnSerThrLysSerLeuLeu 197
QY 424 AATTGTGGAATATGTTAAAGAGAATGCATCAGACATGAACCACTTGCAGAAATCATTTTG 483
Db 198 ProSerLeuIleIleVal-----ProAlaSerLeuLeuLeuAsn 210
QY 484 TGTCGGAACAGTTTATGATTTCTTCAGATATGTGAAATGTCAACATTTGACATAGCT 543
Db 211 TrpLysSerGluPheGlnLysPheAlaProHisValLysIleValThr----- 226
QY 544 TCAGATGTCATTTGCCACATTTCAAGGATTTTACTTACAAGACATAAATGCTCAGTCGAGAA 603
Db 227 -----AlaAsnGlyAsnPheLysGluArgSerGlnValTyrGluSerLeuLysAsnGln 244
QY 604 TTTTGGAAACAGCATTTATGATGATTTTTCAGTGAATAT-----GAGAAAGTTA 651
Db 245 IleLeuLeuMetSerPheAsnValLeuArgSerAspIleLysTrpIleSerGlnLysLys 264
QY 652 CTTTCAT-----TCAGAAATATTGTGACAAAA 678
Db 265 PheHisTyrValValIleAspGluAlaGlnGlyIleLysAsnGluAsnSerThrValThr 284
QY 679 AGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACTTCACAATATG 738
Db 285 LysAlaAlaLysIleLysGlyAsnPheCysLeuAlaLeuThrGlyThrProIleGlu 304
QY 739 ACAAAATACATCAGTAAACCTCAGAACCTCAAAATTAATGATG---AACCTGCTCGGAGAC 795
Db 305 AsnArgLeuLeuAspLeuTrpSerCysPheAppPheValLeuProAsnPheLeuGlyAsn 324
QY 796 AAAAGTCGC---AACATCCAGTTTGTAGGCCTTTACGCTTTTAAAGGTGTTGTAGCCAAT 852

```


Db 1238 LysAspGluAlaIleGlnThrAlaLeuLysGluPheLys----- 1250
QY 427 TGTGGAATAGTAAAGAGAAATGCATCAGACATGAACCACTTCCAAAAATCATTTGTGG 486
Db 1251 -----LeuGluArgGluValGluLysGluLeuLysVal----- 1264
QY 487 TCAGAACAGTTTATGATTTCTCAGATATGTCGAATG----- 525
Db 1265 -----LysHisLeuGluAsnGlnIleAlaLysSerProAla 1276
QY 526 -----TCACATTTGACATGAGCTTCAGATGCAATTTGCCACATCAAGGATTTACTTACA 579
Db 1277 IleAspSerThrArgGlyAspSerSerLeuValAlaGluLeuGlnLysLeuGln 1296
QY 580 AGACATAAATGCTCAGTGCAGAAATTTTGGACACAGCATATCATAGATTTTTCAGTGAA 639
Db 1297 GluLysLys-----AlaLysPheLeuGluGln----- 1308
QY 640 TATGAGAGTTACTTCTATTGAGAAATTTATGTGACAAAAGACAGTCACTGAAGCTTCTC 699
Db 1309 GlnGluLys-----ArgLysAsnGluGluMetGlnAsnValArgThrSerLeuIle 1325
QY 700 GTGAACTACTACTAGATACACAACTTC---ACAATTATGACAAATACATCAGTAA 756
Db 1326 AlaGlu-----GlnGlnThrAsnPheAsnThrValLeuThrArgGluLysMetArg 1342
QY 757 CTGAGAACCTCAAAATTAATGATGAACTGCTGCGAGACAAAAGTCCCAATCCAGTTT 816
Db 1343 LysGluAsnIle-----IleAsnAspSerAspLys----- 1353
QY 817 GAGCCCTTTCAGCTTTTAAAGTGTGTAGCCAACTCTAACAGACGACCCATCCTA 876
Db 1354 -----LeuLysSerThrMetGlnGlnGluArgAspLysAspLeuIle 1368
QY 877 GACATCTCTCAAGAACCCAGGCAAACTCATAGAGTTCTCTCAGCAAGTTTCAGAACGAC 936
Db 1369 GluSerLeuSerGluAspArgAlaArgLeuLeuGluGluLysLysLeuGluGlu 1388
QY 937 AGGACGATTTGATGAGCAGTTCC---GTACCGACGAC----- 971
Db 1389 ValSerLysLeuArgSerSerPheValProSerProTyrValAlaThrAlaProGlu 1408
QY 972 -----GAATCCCGGTCGATTTACGCGTTTAAACCGCGAGCG 1010
Db 1409 LeuTyrGlyAlaCysAlaProGluLeuProGluProGluSerAspArgSerAlaValGluThr 1428
QY 1011 TGGGATCAGGATTTGAAGACACGCTCAGCAAGAAGCTTA 1052
Db 1429 AlaAspGluGlyArgValAspSerAlaMetGluThrSerMet 1442

RESULT 10

US-10-732-923-10741
; Sequence 10741, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10741
; LENGTH: 409
; TYPE: PRP
; ORGANISM: Entamoeba histolytica
US-10-732-923-10741

Alignment Scores:

Pred. No.: 0.0628 Length: 409
Score: 99.50 Matches: 66
Percent Similarity: 33.12% Conservative: 37

Best Local Similarity: 21.22% Mismatches: 111
Query Match: 5.52% Indels: 97
DB: 6 Gaps: 13
US-10-089-688-1 (1-1053) x US-10-732-923-10741 (1-409)
QY 16 GGAAGTCTCAAAATCTCCAGCAGACATTTGGAAGATCTGAAGAGAGCATGGCTGTT 75
Db 2 GlyAsnSerAsnHisLysProGluLysLysAsnAsnLysGlyLysGlyLys 21
QY 76 CTGGAACCAAGACATTTCTGATATAAAAGCAGAAAAG---GCTACAGAAAGATTTCC 132
Db 22 ArgLysAsnGluLysLysLysAspGluThrAsnGluLysGluAlaIleGluProThr 41
QY 133 AAAATCTGTTGCTCCATGAAGAAATTTCTATGGCACAATAATGCAAAAGAGCCTCAGACA 192
Db 42 ProAsn-----GluGluLysGluProCysThr 50
QY 193 GAAGCAGTAGCT-----CACTTGCTCAA----- 216
Db 51 AspThrLysProAsnGluAspAsnGlnIleValLysGlyLysGluGlyLysGluGlu 70
QY 217 -----GAACTCTATAATAGTGGCTCCTTAGCACCTGCTAGCTGATTTACAGCTCAT 270
Db 71 TyrIleGluLeuThrAsnGluGlyLysLysValThrGlnAspAspPheGluLeu 90
QY 271 GACTTTGAGGCG-----AAAAAGACGTGGCT 297
Db 91 LysValIleGlyArgGlySerPheGlyLysValMetValLysLysLysAspAspGly 110
QY 298 CAAATTTTC---AACAAATATCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTGAA 354
Db 111 ArgIlePheAlaMetLysIleLeuArgLysAspIleValLysGluArgLysGlnValAsp 130
QY 355 TACATCTGCACCAACAGAAATATTTGTTTCATGTTTGAAGGGTATGAATCTCCAGAA 414
Db 131 HisThrLysAlaGluLysAsnValLeuMetGlnLeu----- 142
QY 415 ATAGCTCTAAATTTGTGGAATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGA 474
Db 143 -----HisHisProPheIleVal 148
QY 475 ATCATTTTGTGTCGGAACAGTTTATGATTTTCTTCAGATATGTCGAAATGTCACATTT 534
Db 149 LysLeuTyrTyrAlaPheGln----- 155
QY 535 GACATAGCTTCAGATGCTTTGCCACATTCAAGGATTTACTTACAGACATAAATGCTC 594
Db 156 -----ThrThrAspLysLeuTyrMetValMetAspPheValAsnGlyGlyLeu 172
QY 595 AGTCAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAGTTACT 654
Db 173 -----PheTyrHisLeuLysAsnGluAsnCysPheSerGluGluArgAlaLysPhe 189
QY 655 CATTCAAGAAATATGTGACAAAAGACAGCTCACTGAAGCTTCTCGGTGAACTACTACTA 714
Db 190 TyrAlaAlaGluIleAlaThr-----ValLeuIle 199
QY 715 GATAGACACAACTTCACAAATATGACAAAATATACATAGTAACTGAGAACTCAAAATTA 774
Db 200 HisIleHisSerLeuGlyIleIleTyrArgAspLeu---LysProGlu----- 214
QY 775 ATGATGAACCTGCTGCGAGACAAAAGTCGCAACATCCAGTTTGTAGGCGCTTTCAGTTT 834
Db 215 -----AsnIleLeuLeuAspAsnThrGlyAsnIleValIleThrAspPheGlyLeuSer 232
QY 835 AAGGTGTTGTAGGCAATCTTAAACACGCGAG 867
Db 233 LysGlnLeuAlaAlaGly---GluGluThrGln 242

RESULT 11

US-10-732-923-8676
; Sequence 8676, Application US/10732923

```

; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8676
; LENGTH: 1251
; TYPE: PRT
; ORGANISM: Encephalitozoon cuniculi
US-10-732-923-8676

Alignment Scores:
Pred. No.: 0.0926 Length: 1251
Score: 99.00 Matches: 81
Percent Similarity: 36.36% Conservativeness: 59
Best Local Similarity: 21.04% Mismatches: 127
Query Match: 5.49% Indels: 118
DB: 6 Gaps: 20

US-10-089-688-1 (1-1053) x US-10-732-923-8676 (1-1251)

Qy 1 ATGCCGTTCCGTTGGAGTCTCACAATCT---CCAGCAGACATGTGAAGATCTG 57
Db MetThrPheGluGluGlnLysSerGluLysThrArgLysLysAspAlaIleGluThrLeu 819
Qy 58 AAGCAGACATGGCT-----GTTCTGAAAGCAAGACATTTCTGATAAAAGCA 108
Db LeuArgLysGlyAlaTyrglyValLeuMetGluThrAspAlaSerSerArgLysPhe 839
Qy 109 GAAAGGCTACAGAAAGATTTCCAAAATCTGTTGCCATGAAAGAAATCTGTATGGC 168
Db -----CysGluGluAspIleAspGlnIleGluGluArgThrLysIleValLysHis 857
Qy 169 ACAATGAAAAGAGCTTCAGAGAGAGTACTCAACTTGCTCAAGAACTTATAAT 228
Db SerAspGlyGlyAsnValPheSerLysAlaSerPheGlnValGluGluGluAspAsp 877
Qy 229 AGTGGGCTCTAGCACCCCTGGTAGCTGATTACAGCTCATTGACTTTGGGGCAAA--- 285
Db ProAspPheTrpGluAsnLeuSerLysArgLysSerGluGluSerGluGlyArgIle 897
Qy 286 -----AAAGACGTGGCTCAA 300
Db ArgArgGlnMetArgArgLeuAlaArgGluGlyLeuSerSerGluAspIleSerGlu 917
Qy 301 ATTTTCAACATATTTCTCAGAGACAAATTTGGTAGAGAACTCTCTACTGTGTGATATC 360
Db Ile---AspGluLeuLeuAsnThrLysLeuGluSerGluAspGluAsnValPheTy--- 935
Qy 361 TGCACCCACAGATATTTTGTTCATGTTA-----TTGAAA 396
Db ---GluAsnGlnCysLeuLeuIlePheLeuValLeuArgPheGlyProArgSerVal 954
Qy 397 GGGTATGAATCTCCAGAAATAGCTCTAAATTTGTGGAATAATGTTAAGAGAATGCATCAGA 456
Db GlyLeuGluIleProSerLeuThr----- 962
Qy 457 CATGAACCACTT-----GCAAAAATCATTTTGTGGTGGACAGTTT 498
Db HisGluAspLeuLysAspGlyValGluLysAlaHisIleLeuProGlnSerProArgPhe 982
Qy 499 TATGAT-----TTCTTCAGATATGTCGAATATGTCACATTTGAC----- 537
Db LysAspThrAspAlaTyrglyPheArgTyrglyValLysTyrglyValAspGlnPheGlnSer 1002
Qy 538 -----ATAGCTTCAGATGATTGGCCACATTCAGGATTTACTTACAGACATAATTTG 591
Db ThrLysValArgSerAsp-----PheSerAspGlyIle----- 1013

592 CTAGTCAGAAATTTTGGACACACATATGAT-----AGATTTTTCAGT 636
1014 -----GluGluPheLeuValAspAsnTyrglyProSerLeuPheSerLysPheGlySer 1031
637 GAATATGAGAAGTTACTTTCATTCAGAAAATATGTGACAAAAGACAGTCTCAAGCTT 696
1032 ValTyrglyLys-----TyrglyGluArgPheLeuLeuArgIleGlnValValMetIle 1049
697 CTCGGTGAACCTACTACTA----- 714
1050 LeuAsnSerLeuLeuLysThrGluAspLeuAlaValGluLysValArgGlyTrpThrSer 1069
715 -----GATAGACAC-----AACTTCACATATGACAAAATACATCAGTAAACCTGAG 762
1070 AspAspAspArgAlaLeuValSerLeuThrLeuLysTyrglyTyrglyCysTyrgly 1089
763 AACCTC-----AAATTAATGATGAAC 783
1090 ThrIleArgAsnLysSerValAspAspMetAsnGlnArgIleArgLysIleValSerThr 1109
784 CTGCTCGAGACAAAAGTCCCAACATCCAGTTTGAGCCCTTTCACGTTTTTAAGTGTTT 843
1110 LeuSerArgLysLysGluValArgGluAspAspAsnLeuPheHis-----LysAlaIle 1127
844 GTAGCCCATCTTAACAGACGCGAGCCCATCTAGACATCTCTCAAGAACCGGCCAAA 903
1128 MetIlePheGlyArgValThrAsp-----AspAsnArgGluSer 1140
904 CTATAGAGTTCTCTC 918
1141 IleValGluPheLeu 1145

RESULT 12
US-10-732-923-3307
; Sequence 3307, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3307
; LENGTH: 1005
; TYPE: PRT
; ORGANISM: Methanocaldococcus jannaschii
US-10-732-923-3307

Alignment Scores:
Pred. No.: 0.11 Length: 1005
Score: 98.00 Matches: 84
Percent Similarity: 42.99% Conservativeness: 57
Best Local Similarity: 25.61% Mismatches: 120
Query Match: 5.44% Indels: 67
DB: 6 Gaps: 18

US-10-089-688-1 (1-1053) x US-10-732-923-3307 (1-1005)

Qy 49 AAGATCTGAAGAGAGCATGCTGTCTCGAAAGCAA-----GACATT 93
Db LysGluLeuLysAsnLysMetSerGlnLeuGluGluLysAsnLysLysLeuMetGluIle 213
Qy 94 TCTGTATAA-----AAAGCAGAAAAGGTCTACAGAAAGTTTCCAAA-----AATCTG 141
Db AsnAspLysLeuAsnLysIleLysLysGluPheGluAspIleGluLysLeuPheAsnGlu 233
Qy 142 GTTGCATGAAGAAATTTCTGTATGGC-----ACAATGAAAAGAGCTCAGACAGAA 195
Db TrpGluAsnLysLysLeuLeuTyrglyPheIleAsnLysLeuGluGluArgLysArg 253
Qy 196 GCAGTAGCTCAACTTGCTCAAGAACTC-----TATAATAGTGGCTCTTAGC 243
```



```
Db 254 AlaleuGlulLeuLysAsnGlnGluLeuLysIleLeuGluTyrAsp-----LeuAsn 270
Qy 244 ACCGTGTAGCTGATTACAGCTCATTGACCTTTGAGGGCAAAAAGACGGTGGCTCAAAAT 303
Db 271 ThrValValGluAlaArgGlu-ThrLeuAsn---ArgHisLysAspGluTyrGluLysTyr 289
Qy 304 TTCACAAATATTCAGAGAACAAATTTGGTACGAGAACTCTACTGTTGAATACATCTGC 363
Db 289 rLysSerLeuValAspGluLysIleLeuArgLysIleGluSerArgLeu----- 303
Qy 364 ACCCAACAGAAATATTTGTTGTCATGCTATTGAAAGGATGAATCTCCAGAAATAGCTCTA 423
Db 304 -----ArgGluLeuLysSerHisTyrGluAspTyrLeuLysLeuThrLysGlnLeuG1 321
Qy 424 AATTGTGGAATAATGTTAAGAGAAATGATCAGACATCAAGACCACTTGCAAAATCATTTTG 483
Db 321 uileileLys-----GlyAspIleGluLysLeu---LysGluPhe11 334
Qy 484 TGGTCGGAACAGTTTATGATTCTTCAGAT-----ATGTCGAATGTCA 528
Db 334 eAsnLysSerLysTyrArgAspAspIleAspAsnLeuAspThrLeuLeuAsnLysIleLy 354
Qy 529 ACATTTCAGATAGCTTCAGATGCTTTGCCCATTTCAAGGATTTTACTTACAAGACATAAA 588
Db 354 Asp-Gluile-----GluArgValGluThrIleLysAspLeuLeuGluLys 371
Qy 589 TTGCTCAGTCAGAAATTTTGGAA---CAGCATTTATGATAGATTTTTCAGTGA--- 639
Db 372 AsnLeuAsnGluGluileGluLysIleGluLysTyrLysArgIleCysGluGluCysLys 391
Qy 640 -----TATGAGAGTTACTTCATTCAGAAATTTATGTGACAAAAGACATCCTCAAG 693
Db 392 GluTyrTyrGluLysTyrLeuGluLeuGluLysAlaValGluTyrAsnLysLeuThr 411
Qy 694 CTTCGCTGGAATCTACTACTAGATAGACACAACTTCAAAATTTATGACAAAATACATCAGT 753
Db 412 LeuGluTyrIleThrLeuLeuGluGlu-----LysLysSerIleGlu 425
Qy 754 AAACCTGAGAAC---CTCAAAATTAATGATGAACCTGTCGAGACAAAAGTCGCAACATC 810
Db 426 LysAsnIleAsnAspLeuGluThrArgIleAsnLysLeuLeuGluGluThrLysAsnIle 445
Qy 811 CAGTTTGGGCTTTCAGCTTTTAAAGTGTGTTGAGCCATCTCAACAG----- 861
Db 446 AspIleGluSer-----IleGluAsnSerLeuLysGluIleGlu 458
Qy 862 ---ACGACGCCCATCTAGACATCTCTCAAGAACACAGGCCAAACTCATAGAGTTCTCTC 918
Db 459 GluLysLysLysValLeuGluAsnLeuGlnLysGluLysIleGluLeuAsnLysLysLeu 478
Qy 919 AGCAAGTTTCAGAACGCAG 936
Db 479 GlyGluileAsnSerGlu 484
```

RESULT 13

```
US-10-732-923-8707
; Sequence 8707, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8707
; LENGTH: 1359
; TYPE: PR
; ORGANISM: Caenorhabditis elegans
US-10-732-923-8707
```

```
Alignment Scores:
Pred. No.: 0.118 Length: 1359
Score: 98.00 Matches: 82
Percent Similarity: 39.25% Conservative: 64
Best Local Similarity: 22.04% Mismatches: 142
Query Match: 5.44% Indels: 84
DB: 6 Caps: 22
```

US-10-089-688-1 (1-1053) x US-10-732-923-8707 (1-1359)

```
Qy 19 AAGTCTCAAAATCTCCAGACAGCATTTGTGAAGAACTGAAGAGACCATGGCTGTTCTG 78
Db 365 LysLysLysGluLysProGluLysLysLysLysGlyIleIleMetAspSerSerLysLeu 384
Qy 79 GAAAGCAAGACATTTCTGATATAAAAGCAGAAAGGCTACAGAAAGTTCACAAAAT 138
Db 385 GlnLysGluThrIleAsp-----AlaGluArgAlaGluLysGluArgLysArg 401
Qy 139 CTGTTGCCATGAAGAAATTTCTGTATGGC-----ACAAATGAAAAGAGCCTTCAGACA 192
Db 402 LeuGluLysGlnLysGluPheAsnGlyIleValLeuGluGluGluAspLeuThr 421
Qy 193 GAAGCAGTAGCTCAACTGCTCAAGAACTCTATAATAGTGGCTCTTAGCACCTCGTA 252
Db 422 GluMetLeuThrGlyThrSerSerGlnArgLysLysSerValVal-----LeuAsp 439
Qy 253 GCTGATTTACAGCTCATTGACTTTGAGGCAAAAAGACGCTGCTCAAAATTTTCAACAAT 312
Db 440 ProAspSerSerThrValAspGluLysLysProVal-----GluValHisAsnSer 458
Qy 313 ATTCTCAGAGACAAATTTGGTACGAGAACTCTACTGTTGAA-----TACATCTGCACC 366
Db 459 LeuValArgIleLeuLysProHisGlnAlaHisGlyIleGlnPheMetTyrAspCysAla 478
Qy 367 CAACAGAAATTTTGTTCATGTTATTGAAGGGTGAATCTCCAGAAATAGCTCTAAAT 426
Db 479 CysGluSerLeuAspArgLeuAspThrGluGly-----SerGlyGlyIleLeuAlaHis 496
Qy 427 TGT---CGAATA-----ATGTTAAGAGAAATGATCAGACAT 459
Db 497 CysMetGlyLeuGlyLysThrLeuGlnValIleThrPheLeuHisThrValLeuMetHis 516
Qy 460 GAACCACTCTCAAAA-----ATCATTT 480
Db 517 GluLysIleGlyGluLysCysLysArgValLeuValValProLysAsnValIleIle 536
Qy 481 TTGTGTCGGAACAGTTTATGATTTCTTC-----AGATATGTGCAAAATGTCACATTT 534
Db 537 AsnTrpPheLysGluPheGlnLysTrpLeuValAspAsnAspGluGluLeuAspThrIle 556
Qy 535 GACATAGCTTCA---GATGCATTTGCCACATTCAGGATTTACTTACAGACATAAAATTG 591
Db 557 AspValAsnGluLeuAspSerTyrLysThrIleGluAsp-----ArgArgAla 573
Qy 592 CTGAGTCGAGAAATTTTGGAAACAGCATTTATGATAGATTTTTCAGTGAATATGAGAAGTTA 651
Db 574 Leu-----LysAla 576
Qy 652 CTTCATTCAGAAAATTTATGACAAAAGACAGTCACTGAAGCTTCTCGGTGAACACTACTA 711
Db 577 TrpHisSer-----SerLysThrProSerValMetIleIleGlyTyrAspLeu 592
Qy 712 CTAGATAGACACAACTTCACAATTTATGACAAAATACATCAGTAAACCTGAGACCTCAAA 771
Db 593 -----PheArgIleLeuThr---ValGluAspAspProLysLysLysLys 606
Qy 772 TTAATGATGAACCTGCTCGAGACACAAAAGTCCCAACATCTCCAGTTTGTAG---GCCTTTCAC 828
Db 607 -----ProLysAsnArgAsnArgArgLeuGluLysAlaLysGlu 619
Qy 829 GTTTTAAAGTGTGTTAGCCCAATCTTAAACAGACGCGACGCCCATCTTAGACATCTCTC--- 885
Db 620 AspPheArgLysTyrLeuGlnAsnProGlyProAspMetValValCysAspGluAlaHis 639
```

```
QY 886 ---CTCAAGAC---CAGGCCAACTCATAGATTCCTCAGCAGG 939
D 640 LysLeuLysAsnAspSerAlaLeuSerLysCysMetValLysIleLeuThrLysArg 659
QY 940 ACGGATTGTATGACGAGTTCGGTACCGACGACGAAATCCCGGGTCGATTTACGC- 993
D 660 ArgIleCysLeuThrGlyThr---ProLeuGlnAsnLeuMetGluTyrHisCysMet 678
QY 994 -----GTTAAACCGCGACGCGTGGGATCAGG 1020
D 679 ValAsnPheValLysProGlyLeuLeuGlyThrLys 690

RESULT 14
US-10-732-923-8708
; Sequence 8708, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 8708
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-8708

Alignment Scores:
Pred. No.: 0.118 Length: 1359
Score: 98.00 Matches: 82
Percent Similarity: 39.25% Conservative: 64
Best Local Similarity: 22.04% Mismatches: 142
Query Match: 5.44% Indels: 84
D: 6 Gaps: 22

US-10-089-688-1 (1-1053) x US-10-732-923-8708 (1-1359)
QY 19 AAGTCTCAAAATCTCCAGCAGACATTTGTGAAGATCTGAAGAGCAGCATGGCTGTTCTG 78
D 365 LysLysLysGluLysProGlyLysLysLysLysLysLysLysLysLysLysLysLys 384
QY 79 GAAAGCAAGACATTTCTGTATAAAAGCAAGAAAGCGCTACAGAAAGATTTCCAAAAAT 138
D 385 GlnLysGluThrIleAsp-----AlaGluArgAlaGluLysGluArgLysArg 401
QY 139 CTGTTGCCATGAAGAAATTTCTGTATGGC-----ACAAATGAAAGAGCCTCAGACA 192
D 402 LeuGluLysLysGlnLysGluPheAsnGlyIleValLeuGluGluGlyGluAspLeuThr 421
QY 193 GAAGCAGTAGCTCAACTGCTCAGAACTCTATAATAGTGGCTCCTTAGCACCCTGGTA 252
D 422 GluMetLeuThrGlyThrSerSerGlnArgLysLeuLysSerValVal-----LeuAsp 499
QY 253 GCTGATTACAGCTCATTTGAGCGGCAAAAGACGCTGGCTCAAAATTTTCAACAAT 312
D 440 ProAspSerSerThrValAspGluGluSerLysLysProVal---GluValHisAsnSer 458
QY 313 ATTTCTCAGACACAAATTTGGTACGAGAACTCCTACTGTTGAA-----TACATCTGCACC 366
D 459 LeuValArgIleLeuLysProHisGlnAlaHisGlyIleGlnPheMetTyrAspCysAla 478
QY 367 CAACAGAAATATTTGTTTCATGTTATTGAAGGGTATGAATCTCAGAAATAGCTCTAAAT 426
D 479 PheGluSerLeuAspArgLeuAspThrGluGly-----SerGlyGlyIleLeuAlaHis 496
QY 427 TGT---GGAATA-----ATGTTAAGAGAATCATCAGACAT 459
D 497 CysMetGlyLeuGlyLysThrLeuGlnValIleThrPheLeuHisThrValLeuMetHis 516
```

```
QY 460 GRACCACTTCRAAA-----ATCATT 480
D 517 GluLysIleGlyLysCysLysArgValLeuValValProLysAsnValIleIle 536
QY 481 TTGTGTCCGAAACAGATTTTATGATTTCTTC-----AGATATATGCGAAATGTCAACATTT 534
D 537 AsnTrpPheLysGluPheGlnLysTrpLeuValAspAsnAspGluGluLeuAspThrIle 556
QY 535 GACATAGCTTCA---GATGCATTTGCCCATTCACAGATTTTACTTACAGACATAAATTC 591
D 557 AspValAsnGluLeuAspSerTyrLysThrIleGluAsp-----ArgArgAla 573
QY 592 CTCAGTGCAGAAATTTTGGACAGCATTATGATAGATTTTTCAGTGAATATGAGAGTTA 651
D 574 Leu-----LysAla 576
QY 652 CTTTATTTCAGAAATTTATGTGCAAAAAGACACTCACTGAAGCTTCTCGTGAACACTACTA 711
D 577 TrpHisSer-----SerLysThrProSerValMetIleIleGlyTyrAspLeu 592
QY 712 CTAGATAGACAACTTTCACAAATATGACAAATATACATCATAGTAAACCTGAGAACCTCAAA 771
D 593 -----PheArgIleLeuThr---ValGluAspAspProLysLysLysLys 606
QY 772 TTAATCATGAACTGTCGAGACAAAGTCGCAACATCCAGTTTCAG---GCCTTTCAC 828
D 607 -----ProLysAsnArgAsnArgLeuGluLysAlaLysGlu 619
QY 829 GTTTTAAAGTGTTGTAGCCAACTCTAACAGACGAGCCCATCTAGACATCTCTC--- 885
D 620 AspPheArgLysTyrLeuGlnAsnProGlyProAspMetValValCysAspGluAlaHis 639
QY 886 ---CTCAAGAAC---CAGGCCAACTCATAGATTTCTCAGCAAGTTTTCAGAACGACAGG 939
D 640 LysLeuLysAsnAspAspSerAlaLeuSerLysCysMetValLysIleLeuThrLysArg 659
QY 940 ACGGATTGTATGACGAGTTCGGTACCGACGACGAAATTCGCGGTGCTGATTTACGC----- 993
D 660 ArgIleCysLeuThrGlyThr---ProLeuGlnAsnLeuMetGluTyrHisCysMet 678
QY 994 -----GTTAAACCGCGAGCGCTGGGATCAGG 1020
D 679 ValAsnPheValLysProGlyLeuLeuGlyThrLys 690

RESULT 15
US-10-805-684-90
; Sequence 90, Application US/10805684
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748(70342)
; CURRENT APPLICATION NUMBER: US/10/805,684
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,103
; PRIOR FILING DATE: 2003-04-02
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 90
; LENGTH: 1960
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-805-684-90

Alignment Scores:
Pred. No.: 0.145 Length: 1960
Score: 97.50 Matches: 76
Percent Similarity: 36.97% Conservative: 63
Best Local Similarity: 20.21% Mismatches: 166
```

Query Match:	5.41%		71	
DB:	6	Indels:	14	
		Gaps:		
 US-10-089-688-1 (1-1053) x US-10-805-684-90 (1-1960)				
QY	25	CACAAATCTCCAGCAGACATTTGTGAAGAATCTGAAGGAGAGCATGGCTGTTCGAAAAAG	84	
Db	1363	HsAlaGlnValThrAspMetLysIysLysMetGluAaspGlyValGlyCysLeuGluThr	1382	
 QY				
Db	85	CAAGCATTTCTGTATAAAAACGACAAGGCGTACAGAAGAAGTTTCCAAAAATCTCGTT	144	

Search completed: November 10, 2004, 21:15:05
Job time : 45.5 secs